Biomarkers of antidepressant treatment outcomes

Hodgson, Karen

Awarding institution:
King's College London

BIOMARKERS OF ANTIDEPRESSANT TREATMENT OUTCOMES

Karen Hodgson

Student Number: 1050099

Supervisors: Professor Peter McGuffin and Dr Richard J Dobson

August 2014

MRC Social, Genetic and Developmental Psychiatry Centre.

Institute of Psychiatry,

King's College London

Abstract

Whilst antidepressants are widely prescribed, there is a large degree of variation between patients in terms of treatment outcomes. Furthermore, the mechanisms by which these drugs exert their effects remain unclear. In this thesis, genetic biomarkers of antidepressant outcomes have been explored, in order to better understand the molecular mechanisms underpinning effective antidepressant treatment. The research presented here use data from the GENDEP project, which is a large pharmacogenetic study of depressed patients receiving antidepressant treatment.

Firstly, the pharmacological underpinnings of antidepressant-associated side effects were used to categorise these side effects and conduct a candidate gene analysis. Whilst a significant association between variation within the *HTR2C* gene and serotonergic side effects was found, the observation was not replicated in a second sample.

Secondly, the role of variability in drug metabolism rates in treatment outcomes was investigated. Examining genotypic information on the cytochrome P450 enzymes, no associations with treatment response, side effects or study discontinuation were observed. Furthermore, serum concentrations of antidepressant were unrelated to treatment response or overall burden of side effects, predicting only a minority of specific side effects.

Thirdly, transcriptomic changes with drug administration were explored in relation to treatment efficacy. Two genes were identified where changes in expression levels were significantly associated with treatment response amongst patients taking nortriptyline. Furthermore, using a network-based approach, changes in gene expression across one module of coexpressed genes showed significant correlation with symptom improvement; this biological network generalised across different antidepressant medications.

Finally, genomic and transcriptomic data were combined, in an examination of the genetic control of gene expression. This analysis then was used to gain an insight into the molecular processes that link genotype to phenotype.

The evidence presented within this thesis, when considered in combination with existing literature, highlights that antidepressant efficacy is a complex trait, influenced by many genes of small effect. Nevertheless, by layering together different levels of information, we can begin to dissect the molecular mechanisms involved in antidepressant action.

Acknowledgements

Firstly, I would like to thank my supervisors Professor Peter McGuffin and Dr Richard Dobson for all of the support, opportunities and advice given throughout my PhD. I am also grateful for the help I received from Dr Giovanni Coppola whilst at UCLA.

Dr Katherine Tansey and Dr Simone De Jong have been invaluable in their provision of words of wisdom, patience and a good dose of down to earth sanity, and together with many other colleagues at the SGDP have made the last four years a great deal of fun.

Finally, I would like to thank my parents, who have provided endless encouragement, understanding and proof reading services. You can now have your WiFi back, and I promise that this is it; I will finally stop being a student and get a proper job!

Statement of work

The work presented in this thesis was undertaken as part of the clinical component of the Genome-based Therapeutic Drugs for Depression (GENDEP) Project. All statistical preprocessing, analyses and bioinformatics presented here have been performed by K. Hodgson, with two exceptions. For the replication attempt in Chapter 3, Andrew Crawford performed the statistical analysis within the GenPod sample. For the eQTL analysis in Chapter 7, the initial sample quality control applied to the genome-wide data was performed by Rudolf Uher, although the imputation of this data was performed by K. Hodgson. Collection of data, and preparation and laboratory analysis of biological samples were performed by collaborators, as described in the text.

As part of an MSc thesis, K. Hodgson has previously submitted preliminary work on candidate gene analyses of antidepressant side effects in GENDEP. This approach was developed further, and together with replication efforts in an additional sample forms Chapter 3 of this thesis.

The thesis was written by Karen Hodgson, and supervised by Professor Peter McGuffin and Dr Richard Dobson. Karen Hodgson received a studentship from the Medical Research Council (UK) to complete this work.

Table of Contents

Table of Figures

Table of Tables

Chapter 1 Introduction

1.1 Major Depressive Disorder

Major depressive disorder (MDD) is a common and disabling illness; the World Health Organisation has projected that by 2030, it will be the world's number one cause of burden of disease, with the disability burden being 50% higher for females than males (World Health Organisation, 2008). Estimates from the USA suggest that the lifetime prevalence of the disease is 16.6% (Kessler *et al*, 2005).

MDD is characterised by persistent low mood and loss of interest in activities. In order for a patient to fulfil the DSM-IV diagnostic criteria for MDD (American Psychiatric Association, 2000)[1], they should display five (or more) of the symptoms shown in Table 1-1 within the same two-week period, and at least one symptom should be either depressed mood or loss of interest or pleasure.

DSM-IV criteria outline that symptoms should represent a change from previous functioning, and cause significant distress or impairment in social, occupational or other important areas of functioning. Additionally, symptoms should not be due to physiological effects of a substance, a general medical condition or be better accounted for by bereavement.

Table 1-1 Diagnostic symptoms of MDD

Diagnostic symptoms
Depressed mood most of the day or nearly every day
Markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day
Significant weight loss when not dieting or weight gain, or decrease or increase in appetite
Insomnia or hypersomnia nearly every day
Psychomotor agitation or retardation nearly every day
Fatigue or loss of energy nearly every day
Feelings of worthlessness or excessive or inappropriate guilt nearly every day
Diminished ability to think or concentrate, or indecisiveness, nearly every day
Recurrent thoughts of death, suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide

[1] Whilst DSM-V was released in May 2013, the new criteria do not represent a major change from the existing criteria with regards to defining MDD, with the exception of the removal of the "bereavement exclusion". As all studies referred to within this thesis have used DSM-IV criteria, reference is made to this edition of the manual.

Whilst MDD is a single diagnostic entity, patients can differ greatly in their symptomatology, as well as the course and treatment responsiveness of their illness (Rush, 2007). It may be that this symptomatic heterogeneity reflects aetiological and pathophysiological differences between patients.

1.1.1 Aetiology

Genetic factors are known to be important in MDD (Hodgson and McGuffin, 2013),as indicated by evidence from both family studies (Sullivan *et al*, 2000) and newer, molecular genetic methods (Lee *et al*, 2013; Lubke *et al*, 2012). Estimates generally indicate that genetics accounts for somewhere between 30 and 40% of the variability in liability to depression, although there is evidence that more severe forms have higher heritability (McGuffin *et al*, 2007). This genetic risk is thought to comprise of a large number of genetic variants each exerting a small effect, however no specific variants linked to the disorder have yet been identified (Ripke *et al*, 2013). A number of environmental risk factors have been linked to MDD; these include childhood maltreatment (Moskvina *et al*, 2007; Nanni *et al*, 2012), stressful life events (Kendler *et al*, 1999) and neonatal complications (Jablensky *et al*, 2005). Taking the environmental and genetic evidence together, it seems that there is not a single factor which drives the presence or absence of depression. Instead, the disease is best characterised as a complex disorder, where there are likely to be many genes of small effect which act to influence an individual's liability to the illness, as well as a number of environmental factors.

1.1.2 Pathophysiology

In terms of the pathological processes underlying MDD, it has long been considered that dysfunctions in serotonergic signalling play an important role in the disorder (Sharp and Cowen, 2011). But a number of other abnormalities have also been linked to the disorder, including hyperactivity within the hypothalamic-pituitary-adrenal (HPA) axis (Pariante and Lightman, 2008), alterations in neuroplasticity (Pittenger and Duman, 2008) and elevation of proinflammatory cytokines (Dowlati *et al*, 2010). It would seem that as with the disease symptomatology and aetiology, the pathophysiological processes underlying MDD are heterogeneous, with no single unifying abnormality.

1.2 Antidepressants

Despite the limitations in current understanding of the aetiology and pathophysiology underlying MDD, there are a number of treatment options available for depressed patients including pharmacotherapy, psychological treatments, and in more severe cases, electroconvulsive therapy. Antidepressant medications are generally used in moderate to severe cases of depression; there are a number of different drugs available with proven efficacy (Undurraga and Baldessarini, 2012) and these are widely prescribed. For example, in England alone, over 50 million prescriptions were written for antidepressants in 2012 (HSCIC, 2013).

The serendipitous discovery that iproniazid (a monoamine oxidase inhibitor, originally used as a tuberculosis treatment) and imipramine (a tricyclic compound) displayed antidepressant effects led to the development of the first antidepressant treatments during the 1950s. These compounds are thought to exert their effects via the monoaminergic systems, and despite much development and refinement, nearly all of the antidepressants that are currently in use continue to target monoaminergic systems.

Table 1-2: Pharmacology of available antidepressant medications

Drug type	Mode of action	Examples
Tricyclic antidepressants	Inhibition of both noradrenergic and serotonergic uptake (additional action at muscarinic, histaminergic and adrenergic postsynaptic receptors)	Imipramine Nortriptyline
Serotonin reuptake inhibitors	Inhibition of serotonergic reuptake	Escitalopram Fluoxetine
Noradrenaline reuptake inhibitors	Inhibition of noradrenergic reuptake	Reboxetine
Serotonin and noradrenaline reuptake inhibitors	Non-tricyclic compounds which inhibit reuptake of both serotonin and noradrenaline	Venlafaxine Duloxetine
Atypical antidepressants	Various. All show some monoaminergic actions, but not known if these underlie the antidepressant effects.	Agomelatine Mirtazapine
Monoamine oxidase inhibitors		Moclobemide Tranylcypromine

However, despite rapid effects on monoaminergic receptors, there is a significant delay before symptom improvement is observed (Uher *et al*, 2011). This is despite evidence of immediate effects of antidepressants on attention and memory, from both cognitive and imaging studies (Harmer and Cowen, 2013). The evidence indicates that processes occurring downstream from receptor inhibition may be required for clinical efficacy. These mechanisms are yet to be fully elucidated, although many of the systems in which dysfunction is observed in MDD patients appears to be "normalised" with successful antidepressant treatment. For example, it has been noted that whilst patients with MDD are often observed to have low levels of the neurotrophic factor BDNF, antidepressants are observed to increase BDNF levels in MDD patients (Castrén *et al*, 2007). Similarly, abnormalities in the stress-response HPA axis that are seen in MDD are attenuated with antidepressant treatment (Horstmann *et al*, 2009b) and the excess of proinflammatory cytokines often observed in MDD patients are reduced with treatment (Kim *et al*, 2007). Many of these different pathways are being explored for their potential to yield novel-acting antidepressant medications (Berton and Nestler, 2006), but as yet, no single pathway appears to be both necessary and sufficient for therapeutic efficacy.

1.2.1 Efficacy

Indeed, determining the efficacy of these medications in patients has proved a contentious issue. Reports of significant publication bias and small effect sizes relative to placebo amongst patients with milder depression (Khan *et al*, 2002; Kirsch *et al*, 2008) have further complicated the picture (although see Horder *et al*, 2011 for debate surrounding these findings).

Large meta-analyses have attempted to assess the average efficacy of a number of different treatments (Cipriani *et al*, 2009; Gartlehner *et al*, 2011); Gartlehner *et al*. concluded that there is limited evidence of clinically significant differences between medications in terms of response outcomes when looking across a population. Nevertheless, measures of efficacy across a population of depressed patients failed to capture the high variability between patients in terms of treatment response.

This variability was demonstrated within the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study. STAR*D is the largest study of treatment efficacy for MDD to date, including data from over 4,000 patients in the USA. The first line of treatment for patients enrolled in this study was the SSRI citalopram, received for up to 14 weeks; around 30% of patients achieved remission whilst just under 50% of patients showed response to treatment.

Similar findings were also observed in the Genome Based Therapeutic Drugs for Depression study (GENDEP), on which this thesis focuses. GENDEP is a large European pharmacogenetic study, exploring outcomes in patients treated with either the SSRI escitalopram or the tricyclic antidepressant nortriptyline, and for both of these drugs, response and remission rates broadly matched those reported in STAR*D (Uher *et al*, 2009c)

These large studies highlight that antidepressants do work, but only for some. For those that do not respond to the first treatment selected, clinicians must use a trial and error method by which to explore whether other options might be efficacious. With the delay often observed in the onset of treatment response (Uher *et al*, 2011), this can take some time. Inadequate antidepressant treatment results in increased levels of relapse (Kennedy *et al*, 2002) as well as increased health costs (Masand, 2003). Therefore, the development of a rational approach to tailor treatment options to each patient using response-predictive biomarkers would be of significant clinical value. Furthermore, if the biological processes that underlie treatment response can be identified, then it may be possible to use these to guide the development of novel antidepressant medications with improved efficacy.

1.2.2 Tolerability

Antidepressant medications are associated with a number of common side effects, including dry mouth, sexual dysfunction and gastrointestinal symptoms. Given that poor treatment adherence and discontinuation is strongly associated with the experience of side effects (Bull *et al*, 2002; Mitchell, 2006), the tolerability of antidepressant medications is also an important treatment outcome for clinicians to consider. SSRIs are

generally safer in overdose than tricyclic medications (Whyte *et al*, 2003), but side effects are still prevalent with these drugs. Gartlehner *et al* (2011) estimated that in the trials of efficacy amongst second generation antidepressants included in their meta-analysis, 63% of patients experienced at least one antidepressant related side effect. However the side effect profiles do appear to vary between drugs; for example mirtazapine is associated with greater weight gain and paroxetine is associated with more reports of sexual dysfunction that other antidepressants. Similarly, SSRIs are associated with a greater prevalence of diarrhoea and insomnia, whilst tricyclic antidepressants confer a higher risk for dry mouth and constipation (Uher *et al*, 2009a). The observed differences in side effect profiles have generally been observed to vary between drugs in relation to the known receptor actions of the drugs (Stahl, 1998).

Therefore, in addition to the value of identifying treatment response predictors, biomarkers of medication tolerability are also clinically useful. Not only would they indicate which patients may require closer monitoring, but given the differences between drugs in terms of their side effect profiles, if markers of specific ADRs could be identified, then this could be used as an aid to treatment selection.

1.3 Genetics and antidepressant treatment

1.3.1 Pharmacogenetics of treatment response

One approach to identify biomarkers that predict treatment outcomes is to consider the role of genetic factors (Hodgson *et al*, 2012). Clinical observation suggests that response to treatment is a familial trait. Although previously research was constrained by the logistic difficulties in collecting appropriate samples of family members taking the same medications, the sparse familial evidence supported the idea that response to antidepressants was partially genetic in nature (Franchini *et al*, 1998; O'Reilly *et al*, 1994).

But technological advancements mean that it is now possible to use microarray technology to assay genetic variation at upwards of half a million common single nucleotide polymorphisms (SNPs), with relative ease. Using this genomic data, it is possible to directly estimate the proportion of variability observed in a phenotype that is explained by the genetic variability captured by all of the SNPs included on a genomic

microarray (Yang *et al*, 2010). This genome wide complex trait analysis (GCTA) can be done using unrelated individuals, circumventing the requirement to collect twin or family samples in order to estimate the heritability of a trait. Using these methods in antidepressant treatment response, it has been estimated that in a large sample, including subjects from STAR*D and GENDEP, common genetic variants account for 42% of the variation seen between patients in terms of treatment response (Tansey *et al*, 2013). However, GCTA in general may estimate total heritability, and so the overall role of genes in antidepressant response may be even higher.

This genomic data has also been used in genome-wide association studies to try and identify which genetic variants are involved in treatment response. These studies have been able to rule out the presence of single common SNP variants with clinically significant effects of the response outcome (Uher *et al*, 2012), but no replicated association with response has yet been identified. In the largest mega-analysis of antidepressant response to date (n=2,256), no variants reaching genome-wide significance were identified (GENDEP Investigators; MARS Investigators; STAR*D Investigators, 2013), reflecting the limitations in statistical power when conducting analyses to identify small genetic effects on a genome-wide scale (incurring a large multiple hypothesis testing burden) (Pe'er *et al*, 2008). The evidence suggests that whilst response may be under genetic control, it is a complex trait, involving many genes, each with small effects, as well as environmental factors (Tansey *et al*, 2013; Tansey *et al*, 2012). The collection of larger sample sizes is needed to gain sufficient power to reach the threshold for genome-wide significance, but pharmacogenetic cohorts are very expensive to collect (given the need for clinical monitoring across the course of treatment).

Aiming to avoid this multiple hypothesis testing burden, candidate gene studies preselect appropriate genes of interest. These candidates may be either pharmacokinetic (considering “what the body does to the drug”) or pharmacodynamic (addressing “what the drug does to the body”). Although hypothesis-driven methods cannot reveal novel treatment pathways, the approach is better powered to detect genes with smaller effect sizes. However, candidate gene studies have been criticised for poor replicability (Hirschhorn *et al*, 2002; Ioannidis *et al*, 2001). Furthermore, many candidate gene findings suffer from overestimation of effect sizes

in initial studies (known as “winners curse”) (Xiao and Boehnke, 2009). These factors are more likely when sample and effect sizes are small (Ioannidis, 2005), which is frequently the case in candidate gene literature.

Therefore, when interpreting results, key considerations are replication and sample size. There is some evidence from candidate gene work that genes such as the monoaminergic *SLC6A4* and *HTR2A* may predict treatment response, as might the stress-response linked *FKBP5* (Hodgson *et al*, 2012; Horstmann and Binder, 2009a; Kato and Serretti, 2010). However, it should be noted that in genome-wide analyses, these candidates did not show association with treatment outcomes at levels above what would be expected by chance (Garriock *et al*, 2010; Uher *et al*, 2010).

1.3.2 Pharmacogenetics of antidepressant side effects

Genetic predictors of antidepressant side effects have also been explored, although fewer genome-wide studies have been published to date. The largest study comes from STAR*D, where 1,762 patients taking citalopram were considered. The researchers considered a total of five different outcomes; general side effect burden, overall tolerability of treatment, sexual side effects, dizziness and hearing/vision side effects (Adkins *et al*, 2012). In their analysis they found two SNPs reaching significance (defined as an FDR<0.1); hearing/vision side effects were associated with a SNP within *EMID2*, whilst general side effect burden was linked to a gene desert region on chromosome 13. However, these results should be interpreted cautiously, given the number of different outcomes that were considered. Also from the STAR*D sample, but focussing on patients taking drugs other than citalopram, Clark *et al* (2012) performed a genome-wide analysis of four different measures of antidepressant side effects; general side effect burden, sexual side effects, dizziness and vision/hearing side effects. They did not replicate the findings in the larger citalopram-only sample, instead highlighting one association of interest in the subsample of patients taking buproprion (n=128), where SNPs within the *SACM1L* gene were associated with sexual side effects with an FDR<0.05. The final genome-wide study of antidepressant side effects was performed in the GenPod (GENetic and clinical Predictors Of treatment response in Depression) sample of patients taking reboxetine or citalopram, and focussed on sexual dysfunction amongst males (Crawford *et al*, in prep). However, no associations reached genome-wide significance.

As in the treatment response literature, these studies of antidepressant side effects require further replication and are limited by constraints of sample size. However, an additional barrier to identifying genetic associations with antidepressant side effects is the number of different side effect outcomes that can be considered.

This can also be seen when considering the candidate gene research into genetic predictors of side effects (for review, see Kato *et al*, 2010). A wide range of different outcomes are defined, with some reports considering overall side effect burden (Hu *et al*, 2007; Murphy *et al*, 2004), whilst others focus on specific adverse drug reactions (ADRs) such as sexual dysfunction (Perlis *et al*, 2009; Strohmaier *et al*, 2011) or nausea (Sugai *et al*, 2006; Tanaka *et al*, 2008). But a diffuse literature with frequently small sample sizes and poor coverage of so many different outcomes means that no variants have been robustly associated with adverse reactions.

The outcome that has been most frequently reported on is that of overall side effect burden. Prediction of this may be clinically useful, identifying patients who require closer monitoring and careful dosing. However, this approach cannot be used to direct decisions as to which antidepressant might be most appropriate for a patient. Given the different side effect profiles of different antidepressants, information of risk for specific ADRs may be more helpful in this respect.

There are a large number of different side effects to consider, but these ADRs often share a common pharmacological basis. For example, dry mouth, blurred vision, constipation and problems with urination are all classic anticholinergic side effects, found with many drugs that have an antagonistic action at muscarinic acetyl choline receptors. Whilst tricyclic antidepressants show anticholinergic action, SSRIs do not. Thus identifiers of risk for specific ADRs which can be tied to the pharmacological actions of specific medications may be useful in guiding treatment options.

1.3.3 Chapter 3 aims; identification of genetic predictors of antidepressant side effects

Given the large number of different side effects reported with antidepressant treatment, in Chapter 3 of this thesis the shared pharmacological basis of ADRs will be used to rationally reduce the number of different side effect outcomes. Then a candidate gene analysis will be performed to identify predictors of antidepressant side effects in a manner that aims to be useful in guiding treatment recommendations.

1.4 Cytochrome P450 enzymes and antidepressant treatment

As noted above, one of the constraints when using genome-wide approaches is the limitation in statistical power, due to a large multiple hypothesis testing burden. In addition, these approaches consider only individual SNP effects in an independent fashion. This means that other forms of genetic variation such as gene duplications will not be captured. Furthermore, variants with common effects (for example three different SNPs all resulting in a non-functional protein) will be tested separately. This is of particular importance when considering the genetic variation seen in the cytochrome P450 enzymes.

1.4.1 Genetic variation in cytochrome P450 enzymes

The cytochrome P450 (CYP450) enzymes are members of an isoenzyme superfamily that catalyse the oxidation of a large number of different chemicals, including the majority of prescribed drugs. Therefore, they are central in the pharmacokinetics of many medications. Many of the CYP450 enzymes are highly polymorphic, for example, over 100 allelic variants have been described for CYP2D6, and more than 30 allelic variants have been identified within CYP2C19. Variants including splicing defects, missense mutations, frameshifts and duplications have been observed, with functional effects on levels of enzyme activity. Duplications and variants increasing expression levels have been linked to increased enzymatic activity whilst other variants decrease function, or in the case of a null allele lead to the enzyme not being encoded. The different variants that have been described are catalogued by the Human CYP450 Allele Nomenclature Committee (http://www.cypalleles.ki.se/), with any functional effects detailed.

Given the number of different polymorphisms observed, scoring systems have been devised to translate these variants into categories with common functional effects. Whilst the definitions vary, in general, individuals who carry two null alleles are designated as poor metabolisers. Intermediate metabolisers are those who carry one null allele with one functional allele, or who carry alleles with reduced function. Extensive metabolisers have two functional alleles; individuals in this category are considered to have normal enzyme activity. Finally, individuals who carry gene duplications are designated as ultra-metabolisers, with higher rates of drug metabolism.

Whilst GWAS microarray chips only directly capture a small number of the SNPs observed in the CYP450 enzymes, microarrays have been developed which specifically test the known functional variants (including gene duplications) that are reported in these genes (for example the Affymetrix Drug Metabolising Enzymes and Transporters Panel, and the Roche AmpliChip CYP450 Test). Algorithms to derive appropriate functional groupings to this genetic data are then applied.

In 2004, FDA approval was granted for the Roche Amplichip CYP450 test for use by physicians (de Leon *et al*, 2006). This has sparked a growing interest in the potential role CYP450 genes might play in informing clinical practice for a number of different conditions. Indeed, the majority of health insurers in the USA currently cover provision of the test. However, as Matchar and Thankur highlight, FDA approval is linked to the accuracy of the technology in providing genotypic information, not the demonstration of its impact on clinical outcomes (Matchar and Thakur, 2007a). Nevertheless, CYP450 enzymes are of interest when considering outcomes with antidepressant treatment, given their role in the metabolism of these drugs. The particular CYP450 enzymes and metabolic pathways involved vary between the drugs, and thus must be considered in a drug-specific manner.

1.4.2 Metabolism of escitalopram

Escitalopram is rapidly absorbed following oral administration, with maximum plasma concentration typically reached after 4 hours. The drug is predominantly metabolised in the liver, with an elimination half-life of 27-

32 hours (Gutierrez and Mengel, 2002). The drug is demethylated to its primary metabolite des-methylcitalopram. CYP2C19 is the key enzyme involved in this demethylation pathway, but CYP3A4 and CYP2D6 also play a role. Escitalopram is a weak inhibitor of CYP450 enzymes, so conferring minimal risk of drug interactions (Gutierrez *et al*; von Moltke *et al*, 2001). Des-methylcitalopram is present in much lower concentrations, and displays only weak inhibition of the serotonergic transporter in comparison to the parent drug (von Moltke *et al*, 2001).

In terms of genetic variability impacting on the function of the CYP450 enzymes involved in the metabolism of escitalopram, functional effects of common polymorphisms have been observed for both *CYP2C19* and *CYP2D6* (de Morais *et al*, 1994; Eichelbaum *et al*, 1979; Ferguson *et al*, 1998). However, genetic variation in the *CYP3A4* gene has been reported to be rare, with little impact on enzymatic activity (Lamba *et al*, 2002).

1.4.3 Metabolism of nortriptyline

For nortriptyline, peak plasma concentrations occur around 7-8.5 hours after oral administration. It is also metabolised in the liver, and has an elimination half-life of around 32 hours. CYP2D6 is the key enzyme involved in the metabolism of nortriptyline; it is responsible for around 90% of metabolism (Olesen and Linnet, 1997a). The primary metabolic pathway is via the CYP2D6 mediated hydroxylation to E-10-hydroxynortriptyline. However, a secondary, minor pathway involves the demethylation of the drug to form des-methylnortriptyline. CYP2D6 is also the most important enzyme in this secondary pathway. Nortriptyline shows weak inhibition of CYP2D6 and CYP2C19, which has been shown to be clinically insignificant when considering drug interactions (Gillman, 2007). The metabolites of nortriptyline show only weak noradrenergic action (Nordin and Bertilsson, 1995).

1.4.4 CYP450 genotypes and serum concentration of drug

Linking together what is known about the functional impact of polymorphisms within *CYP2D6* and *CYP2C19* with their importance in the metabolism of antidepressants, it has been shown that the differences observed

between patients in serum concentration of antidepressant (Hammer and Sjoqvist, 1967; Reis *et al*, 2007) are partially determined by genetic variability in the CYP450 enzymes (Kirchheiner *et al*, 2004).

In a large systematic analysis of the relationship between CYP450 genotype and serum concentration of psychiatric medications, dose adjustments which would compensate for the observed differences in drug metabolism were proposed. For patients taking nortriptyline with PM status for the *CYP2D6* gene, a dose reduction of 53% of the average dose was suggested. When considering citalopram (data on escitalopram was not available), a dose adjustment of 61% was calculated for individuals with poor metaboliser status for *CYP2C19*. Since this report, additional studies addressing the relationship between CYP450 genotype and serum levels of antidepressant have been published looking at escitalopram (including one study looking within a subset of the GENDEP sample); these observe a similar relationship between *CYP2C19* genotype and serum concentrations of escitalopram (Huezo-Diaz *et al*, 2012; Rudberg *et al*, 2006; Rudberg *et al*, 2008).

1.4.5 CYP450 genotypes and treatment outcomes

Given the observed variability in antidepressant metabolism via the CYP450 enzymes, there has been much hope regarding the potential to use genetic variability in these enzymes to predict outcomes and guide treatment recommendations (Ingelman-Sundberg, 2004; Kirchheiner and Rodriguez-Antona, 2009). This is predicated on the assumption that if individuals who metabolise the drugs very rapidly have lower concentrations of drug, this may lead to insufficient levels of antidepressant at the site of drug action, causing treatment non-response. Conversely, if individuals with low rates of drug metabolism have high levels of antidepressant at the site of drug action, this will cause higher risk of adverse drug reactions.

1.4.5.1 CYP450 genotypes and treatment outcomes for SSRIs

However, in an Evidence Report from the Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group (EGAPP Working Group, 2007; Matchar *et al*, 2007b), the researchers concluded that the limited existing data exploring the link between SSRIs and either treatment response or treatment

tolerability was limited by small sample sizes and heterogeneous study design. They concluded there was no evidence of an association between CYP450 genotype and treatment outcomes for these antidepressants.

Several larger studies have been published since the EGAPP Working group evaluation, but these generally support the conclusions that CYP450 genotype is not linked to treatment outcomes. The largest of these studies is STAR*D; in a two-stage analysis of 1,953 patients taking citalopram, Peters *et al* found no association between any of the pharmacokinetic genes they considered (*CYP2D6, ABCB1, CYP2C19, CYP3A4*, and *CYP3A5*) and either treatment response or tolerability (Peters *et al*, 2008). Given the ethnic variation in the STAR*D sample, this analysis was conducted separately in Caucasians and African American sub-groups. Whilst a second paper has been published on the same sample, concluding that *CYP2C19* genotype was associated with both remission and tolerability when the Caucasian analysis was limited to non-Hispanics (Mrazek *et al*, 2011), the associations fail to reach significance when correcting for the number of hypotheses tested.

Supporting the conclusions of Peters *et al*, no association between CYP450 genotype and response was observed in a sample of 278 patients taking a range of different antidepressants (Serretti *et al*, 2009), nor between CYP450 genotype and either response or side effects in a sample of 106 patients taking either escitalopram or venlafaxine (Ng *et al*, 2013). In contrast, Tsai *et al* did observe a link between *CYP2D6* and treatment response in 100 patients taking escitalopram. However, in this case, genotype was not linked to serum levels of drug, and *CYP2D6* only plays a minor role in the metabolism of escitalopram. When the primary enzyme of *CYP2C19* was considered, associations were seen with serum concentrations of drug but not response.

Therefore, whilst there are some discrepancies between findings, more recent findings appear to support the conclusions of the EGAPP, and there is an absence of strong evidence linking CYP450 genotype to antidepressant response or side effects for patients taking SSRIs.

1.4.5.2 CYP450 genotypes and treatment outcomes for tricyclic antidepressants

In contrast, a recent review from the Clinical Pharmacogenetics Implementation Consortium, looking at tricyclic antidepressants, gives strong recommendations to avoid prescribing either amitriptyline or nortriptyline to individuals identified as *CYP2D6* ultra metabolisers or poor metabolisers. Amitriptyline is metabolised to nortriptyline via CYP2C19 (Olesen and Linnet, 1997b). The authors of the report suggested ultra-metabolisers are at risk of lack of efficacy, whilst poor metabolisers are likely to suffer side effects. However, when reviewing the evidence that is presented as part of their systematic literature review, the guidelines reference only three studies which observe an association between drug metabolising variables and side effects to antidepressant treatment, with no association made to treatment response.

In the first study, a small sample of 18 patients taking a range of different antidepressants, side effects were noted to be more common amongst patients with alleles conferring low CYP2D6 activity (Chen *et al*, 1996). More convincingly, in a sample of 50 patients taking a fixed dose of amitriptyline, nortriptyline plasma concentrations and *CYP2D6* genotype were found to be associated with side effects, but not treatment response (Steimer *et al*, 2005). Similarly in the Rotterdam study (Bijl *et al*, 2008) (with a large total sample size of 1,198) there was increased risk of antidepressant switching amongst patients taking tricyclic antidepressants with poor metaboliser genotypes. However, no association was seen with treatment discontinuation in this sample. The authors also found that amongst patients taking SSRIs, no association was observed between *CYP2D6* genotype and any of the clinical outcomes considered.

Thus, there is some evidence indicating CYP450 genotype might be linked to the tolerability of tricyclic antidepressants, but no evidence of an association with treatment response.

1.4.6 Serum concentrations of antidepressant and treatment outcomes

Whilst the existing literature suggests CYP450 genotypes do play an important role in antidepressant metabolism rates, the difficulty in linking this to treatment outcomes may reflect that drug metabolism is also influenced by other, environmental factors. These include diet, comedications and comorbidities (Reis 2009). By directly measuring serum concentrations of drug, the influence of both genetics and the environment can be considered together.

But the evidence linking serum concentrations to outcomes with antidepressant treatment is also mixed. A seminal study in 1971 reported a curvilinear relationship between serum concentrations and treatment response amongst patients taking nortriptyline (Asberg *et al*, 1971), and this relationship was also observed more recently (Perry *et al*, 1994), with therapeutic concentrations estimated to be between 50-150 ng/ml. Nevertheless, determining therapeutic ranges for antidepressants can be imprecise due to confounds of spontaneous remission, placebo response and the heterogeneity of the disorder (Preskorn and Fast, 1991), and other samples have reported no evidence linking serum levels of nortriptyline to response (Steimer *et al*, 2005). For SSRIs, therapeutic ranges are less clearly defined (Baumann, 1996) and whilst one study suggested that for patients taking citalopram, serum concentrations below 50ng/mL were associated with reduced response (Ostad Haji *et al*, 2011), several other papers indicate there is no link between serum levels of antidepressant and response (Dufour *et al*, 1987; Nikisch *et al*, 2004; Rasmussen and Brosen, 2000).

1.4.7 Chapter 4 and 5 aims; the role of drug metabolism variables in treatment outcomes

Given the interest surrounding the potential role of cytochrome P450 enzymes in determining outcomes with antidepressant treatment, but the absence of robust evidence, in Chapters 4 and 5, I aim to assess their impact in the GENDEP sample. Using both genotypic, serum concentration and treatment outcome data, the interrelationship between genetic variability, drug metabolism rates and antidepressant response, side effects and drop out will be assessed.

1.5 Gene expression changes and mechanisms of antidepressant action

As detailed above, the mechanisms by which antidepressants exert their therapeutic effects remain unclear. As researchers have looked at the biological processes affected beyond the immediate effects of the drugs on neurotransmitter levels and receptors, there has been an increasing focus on gene expression as a possible process which may underpin the long term adaptations in neuronal function that appear necessary for effective treatment (Duman *et al*, 1997; Lesch and Schmitt, 2002).

In light of the high inter-individual variability seen in antidepressant treatment response, it is important not to simply identify changes that occur as a result of the administration of a drug. To uncover the mechanisms by which antidepressants exert their effects, those changes that are specifically associated with response are of interest.

1.5.1 Tissue-specificity of gene expression

The primary tissue of interest when considering antidepressant action is the brain. However, brain samples from patients are evidently only obtainable from post-mortem tissue; this means gene expression can only be studied at a single time point. Therefore, in order to consider the change in gene expression across time, and its association with response, blood samples have been used instead. Blood has been shown to be a useful surrogate for investigating gene expression; for example one study reports that blood shares more than 80% of the transcriptome with nine tissues including brain (Liew *et al*, 2006). Furthermore, comparisons between prefrontal cortex and blood indicate that those processes identified as potentially important in both the pathophysiology of MDD and the mechanisms of antidepressant action (such as neurotransmitter, stress mediator and cytokine pathways) are also overlapping between the two tissues (Sullivan *et al*, 2006).

1.5.2 Changes in expression levels of candidate genes

When exploring gene expression changes associated with antidepressant response, the majority of studies to date have used candidate gene approaches. A number of different associations with response outcomes have been reported. Monoaminergic systems have been implicated in a study of 11 MDD patients, where

changes in expression levels of *SLC6A4* (the gene for the serotonin transporter) were correlated with response to a range of different forms of treatment (Belzeaux *et al*, 2010). Reports highlighting neurotrophic gene expression changes include an observed correlation between *VEGF* (the gene encoding vascular endothelial growth factor) and symptom improvement in a sample of 24 patients (Iga *et al*, 2007a), as well as differential expression between responders and non-responders for the genes *BDNF* and *VGF* in a subset of 74 patients from the GENDEP sample (Cattaneo *et al*, 2013). This study also implicated both the HPA-axis linked *FBKP5* and the inflammatory gene *IL-6* as treatment-correlated genes. In a second study exploring an independent set of 46 patients from GENDEP, a second association between response and the inflammation-linked gene *ABCF1* (ATP-Binding Cassette, Sub-Family F, member 1) was observed (Powell *et al*, 2013).

Whilst these findings are consistent with our growing knowledge regarding the role of monoaminergic, neurotrophic, stress response and inflammatory pathways in antidepressant action, it should be noted that many of these studies are small in size, and are currently without replication. Exacerbating this issue, several of the studies above have selected different subsets of genes within the same putative pathways of interest.

1.5.3 Transcriptomic methodologies

Tackling this issue, it is now possible to move beyond candidate gene expression studies to transcriptomic approaches. These have the advantage of offering systematic assessment of all genes, and as they are hypothesis-free, can also highlight novel associations. However, the advantages of this comprehensive approach must be balanced against the large multiple hypothesis testing burden, reducing the power to detect effects amongst small samples.

Two small studies (n<10) have been published comparing expression levels before and after treatment on a transcriptome-wide scale, but without reference to treatment efficacy. Whilst in one report, no significant transcriptomic expression changes are associated with antidepressant treatment (Belzeaux *et al*, 2012), in the second study changes within genes related to ionic homeostasis and neuronal plasticity were highlighted

(Kálmán *et al*, 2005). However, as no multiple hypothesis testing correction was applied, these associations are likely to be false positives.

The largest study to date exploring transcriptomic changes with antidepressant treatment used data collected from 63 MDD patients (Mamdani *et al*, 2011). These patients were all treated with citalopram, and the authors observed that the gene *IRF7* showed change in expression levels that was differentially linked to treatment response. Furthermore, additional analysis in post-mortem brain tissue also indicated that in prefrontal regions expression of this gene was decreased in MDD.

1.5.3.1 Networks of gene coexpression

The transcriptomic studies detailed above consider each gene and its association with treatment response independently. However, the regulation of gene expression acts via co-ordinated networks; structured correlation patterns between genes are observed (Lee *et al*, 2004; Oldham *et al*, 2008). In order to capture the interconnected nature of gene expression, network based approaches to transcriptomic data have been developed (Langfelder and Horvath, 2008). Modules of genes which share coexpression patterns are identified, and these co-regulated modules are likely to be functionally related. By considering the transcriptomic data using this systems-based level of analysis, the interconnected nature of gene expression is captured and any trait-associated signals are aggregated over a functionally-related set of genes. Furthermore, these network-based approaches have the added benefit of reducing the multiple hypothesis testing burden and so increasing statistical power to detect effects. The methodology has been successfully applied to a number of phenotypes including schizophrenia (de Jong *et al*, 2012), cancer (Horvath *et al*, 2006) and Alzheimer's disease (Miller *et al*, 2008), but has not yet been explored in relation to antidepressant treatment response.

1.5.4 Chapter 6 aims; identification of gene expression correlates of treatment response

In light of the limited literature on transcriptomic correlates of treatment response, in Chapter 6 of this thesis, I aim to explore gene expression correlates of antidepressant response on a transcriptomic scale, within a

subset of the GENDEP sample. This will not only be the largest sample studied using transcriptomic data to date, but also the first example in which network based methodologies have been explored.

1.6 Genetic control of gene expression to understand predictors of response

1.6.1 Discovery of genetic associations and linking these to phenotypes

Attempts to disentangle the genetics of antidepressant response have been discussed above, and since the publication of the human genome sequence in 2001, there has been an ever growing number of genetic variants which have been linked to individual differences in a huge variety of different phenotypes (see NHGRI GWAS catalogue; http://www.genome.gov/gwastudies/). However, as the number of phenotype-associated variants grows, there is an increasing need to understand how these variants are linked to the phenotype of interest. Whilst initial expectations were that variants would be within gene-coding regions and take the form of non-synonymous or missense mutations, 90% of trait-associated variants are outside of protein coding regions of the genome (Hindorff *et al*, 2009). Therefore the manner in which they exert their effects is unclear and increasingly, researchers are looking at subtler mechanisms by which genotype might impact on biology.

1.6.2 Expression quantitative trait loci

Jansen and Nap published an article in 2001 which discussed the potential power of combining genetic and transcriptomic data (within a field they termed “genetical genomics”) in order to trace the impact of genetic variants on biological pathways (Jansen and Nap, 2001). Since then, a growing number of studies have systematically explored the genetics of gene expression. As gene expression levels are measured as a quantitative trait, the genetic loci which influence expression levels are referred to as expression quantitative trait loci, or eQTLs. Similar approached have also been developed to consider genetic drivers of methylation (Gibbs *et al*, 2010), protein (Lourdusamy *et al*, 2012) and metabolite levels (Nicholson *et al*, 2011), each giving information of different levels along the pathway between genotype and phenotype.

Using transcriptome-wide approaches it has been shown that genetic influences on expression levels are wide-spread throughout the transcriptome (Lappalainen *et al*, 2013). These genetic influences are generally classified into two groups; those which are *cis*-acting, where the variant is local to the affected gene, and those which are *trans*-acting, where the variant is distant to the affected gene (potentially located on a different chromosome). By building a resource of well characterised eQTL loci, it is possible to gain a handle on the impact of genetic variation on expression levels, the relative importance of *cis*-eQTLs compared with *trans*-eQTLs, and to begin to understand how trait-associated genetic variants may exert their effects.

Studies have more frequently focussed on *cis*-eQTLs, and these are generally observed to be located near to the transcription start site of a gene, or within exonic regions of the gene (Dimas *et al*, 2009; Stranger *et al*, 2007; Veyrieras *et al*, 2008). Furthermore, *cis*-eQTL signals are often observed to be overlapping with activating cis-regulatory elements including transcription factor binding sites (Gaffney *et al*, 2012) and DNase-1 hypersensitive sites (Degner *et al*, 2012)

The effect sizes of *trans*-eQTLs are generally smaller than those observed for *cis*-eQTLs (Grundberg *et al*, 2012). Nevertheless, *trans* effects are significant, and thought to be more numerous than *cis* effects (Schadt *et al*, 2003), with a number of posible “master regulators” identified, where genetic variation controls expression levels of many genes (Morley *et al*, 2004). *Trans*-eQTLs are of particular interest as they reveal information about the connection between physically distant genetic and transcriptomic variants based on biological consequences. Nevertheless, eQTL analysis is still a rapidly developing field; as more data is collected, understanding of *cis* and *trans* effects on gene expression will grow.

1.6.3 eQTLs as a bioinformatic tool

1.6.3.1 Using eQTLs to interpret genetic and transcriptomic associations

Once eQTL loci have been identified, they can be utilised as an important bioinformatic annotation. There are a growing number of examples where eQTL annotations have been used to aid the interpretation of genetic signals by identifying common downstream effects on gene expression (Fehrmann *et al*, 2011;

Mäkinen *et al*, 2014), or understand transcriptomic associations, where disturbances in expression levels of a number of genes are shown to be driven by genetic variation at a single *trans*-acting SNP (Small *et al*, 2011; Westra *et al*, 2013).

1.6.3.2 Using eQTLs to identify genetic associations

In addition to their value in interpretation, eQTL annotations may also aid the identification of trait-associated genes. Underscoring the importance of the regulation of gene expression, it has been noted that there is an enrichment of eQTLs amongst trait-associated GWAS signals (Nicolae *et al*, 2010). This enrichment can be exploited using a Bayesian approach to up-weight known eQTLs (and so prioritise these variants) prior to a genome-wide analysis, and so increase the ability to detect true signals within the data (Knight *et al*, 2011).

In areas such as antidepressant treatment response, where genome-wide associations are yet to be found, and the collection of samples is very expensive, this overrepresentation of eQTLs amongst trait-associated SNPs may be useful, providing a method by which to recover genomic signal within the noise of GWAS data.

1.6.4 Context specificity of eQTLs

While many eQTL analyses are performed in blood samples taken from control subjects, due to ease of access, one caveat to consider when using eQTL annotations is that the regulation of expression levels by genotype may vary depending upon context.

For example, as gene expression profiles are known to vary between tissues, so do patterns of eQTLs. To explore the tissue-specificity of eQTLs, both the MuTHER project (Multiple Tissue Human Expression Resource; http://www.muther.ac.uk/) and the Genotype-Tissue Expression project (GTEx; http://www.gtexportal.org/) have been set up to characterise eQTLs in a range of tissues. Currently available estimates of the degree of eQTL overlap between tissues vary according to both the analysis approach used and the tissues that are compared. For example, a pilot study from the MuTHER study estimated that 30% of

cis-eQTLs were shared between the tissues studied (Nica *et al*, 2011), but when the sample was expanded, 60% of identified *cis*-eQTLs were found to have significant effects in multiple tissues (Grundberg *et al*, 2012).

The issue of tissue-specificity indicates that ideally, when probing gene-trait relationships, analyses should focus on eQTLs that have been identified in the primary tissue of interest for that trait. For example, adipose tissue samples were used in the identification of a *trans*-regulator of gene expression linked to diabetes (Small *et al*, 2011). However, eQTLs identified in blood have been informative as to the downstream effects of trait-associated SNPs in non-blood related traits such as breast cancer and ulcerative colitis (Fehrmann *et al*, 2011). Furthermore, in the case of antidepressant treatment response, whilst the brain may be the primary tissue of interest, differences in eQTLs have been observed between brain regions (Hernandez *et al*, 2012; McKenzie *et al*, 2014), and the region that would be relevant for treatment response is unclear.

Phenotypic specificity

There is evidence that genetic control of gene expression varies not only with tissue type, but also with other factors. For example sex and age-specific eQTLs have been noted (Dimas *et al*, 2012; Glass *et al*, 2013; Yao *et al*, 2014). Context-specific eQTLs have been robustly indicated using an *in vitro* paradigm of immune-stimulating cells using interferon-γ or lipopolysaccharide (Fairfax *et al*, 2014). Over 50% of the *cis*-eQTLs identified were only observed in specific states of immune activation, and a number of context-specific *trans*-eQTL hotspots were also noted. Interestingly, these hotspots were within immune-related regions of the genome. Thus, the presence of eQTLs within immune-associated pathways is dependent on the immune context of the tissue.

The characterisation of context-specific eQTLs is still in its infancy, and further work is needed to establish its importance. However, given that gene expression differences have been shown to predict antidepressant response in depressed patients (Mamdani *et al*, 2013; Tansey *et al*, in prep), it may also be the case that the

genetic regulation of gene expression levels varies between patients who are more or less likely to respond to antidepressant treatment.

If antidepressant response-specific eQTLs can be identified, this indicates that the mechanism by which a genetic variant influences expression levels varies in a manner which is associated with the treatment response phenotype, and so gives an insight into the biological differences that may exist between responders and non-responders prior to treatment.

1.6.5 Chapter 7 aims; identification and exploration of eQTLs

In chapter 7 of the thesis, I aim to use both the genetic and transcriptomic data available to characterise the eQTLs that can be observed within the GENDEP dataset. Having identified the pattern of eQTLs, I will investigate their potential value as bioinformatics annotations in identifying predictors of antidepressant treatment response. Finally, I will consider whether the genetic effect of gene expression levels is dependent on the phenotype of antidepressant treatment response.

1.7 Conclusions

Given the widespread use of antidepressants, there is a need to better understand and predict the high degree of variability seen in treatment outcomes. Genetics offers one tool by which to approach this issue, with the advantage that only a blood sample is required from the patients. However genetic approaches often come at the cost of low statistical power, requiring the collaborative collection of very large samples within projects such as GENDEP. In this thesis, I will use a range of different genetic-based approaches applied to the GENDEP sample, to attempt to disentangle the biological underpinnings of effective antidepressant treatment, with the ultimate aims of informing clinical decision making and identifying novel targets for antidepressant drug development.

Chapter 2 Methods

2.1 The GENDEP project

This thesis uses data collected as part of the Genome-based Therapeutic Drugs for Depression (GENDEP) project. The project includes *in vitro* and animal components, but the work here comes solely from the human pharmacogenetic portion of the project. GENDEP was approved by ethics boards in each of the participating centres and all participants provided written consent after the procedures were explained. GENDEP is registered at EudraCT (No.2004-001723-38, http://eudract.emea.europa.eu) and ISRCTN (No. 03693000, http://www.controlled-trials.com). The project was designed to identify clinical, genetic and environmental predictors of antidepressant treatment response. Below, the measures that were used within this thesis are described.

2.2 Participants

Participants were recruited into the study across nine centres within eight European countries (detailed in Table 2-1), using generalist and specialist referrals as well as advertisements. For inclusion in the study participants must have been diagnosed with at least moderate severity of depression, as defined by ICD-10 (World Health Organisation, 1992) and DSM-IV (American Psychiatric Association, 2000). Diagnosis was established using the semi-structured SCAN interview (Wing *et al*, 1998), and the computerised classification system CATEGO5 (Grayson *et al*, 1990). Exclusion criteria included; a first-degree relative with bipolar disorder or schizophrenia, a history of hypomanic or manic episodes, mood incongruent psychotic symptoms, and current dependence on drugs or alcohol. Patients who had contraindications (including previous history of lack of efficacy or side effects) to both of the study medications were also excluded, as well as those who were either pregnant or lactating. All patients were of White European descent, in order to control for population stratification within genetic analyses, and over 18 years of age. The complete sample totalled 868 patients; 547 were female and 321 were male. The age range of the sample was 19 to 72 years (Mean = 42.56 years, SD = 11.71).

Table 2-1: Number of patients recruited into GENDEP per centre

Centre	Number of patients recruited
London	109
Brussels	38
Mannheim	89
Bonn	132
Brescia	46
Aarhus	92
Ljubljana	109
Poznan	107
Zagreb	146

2.3 Antidepressants

Two antidepressant drugs were used in GENDEP; escitalopram and nortriptyline. These drugs were selected because they differentially target the two most common mechanisms of antidepressant action. Whilst escitalopram is an SSRI which acts via inhibition of serotonin reuptake, nortriptyline is a tricyclic antidepressant which predominantly influences noradrenaline reuptake action. More specifically, escitalopram has a very high affinity for the serotonin transporter and negligible affinity for other receptors in the brain (Sanchez, Bergqvist et al. 2003). In contrast, nortriptyline shows preferential action at the noradrenergic transporter, although the drug displays some affinity for the serotonin transporter (Sanchez and Hyttel 1999). However, nortriptyline also demonstrates antagonistic action at several postsynaptic receptors; the H1 histamine receptor, the adrenergic α-1 receptor, serotonergic 5HT-2A and -2C receptors and muscarinic cholinergic receptors (Sanchez and Hyttel 1999). Whilst reboxetine may be more selective in its action at noradrenergic receptors (Wong *et al*, 2000), there are questions regarding its efficacy (Eyding *et al*, 2010), whilst both escitalopram and nortriptyline have good efficacy records (Cipriani *et al*, 2009; Nelson, 1999).

Patients were randomly allocated to receive one of these two antidepressants, unless prior contraindications indicated one of the drugs was not suitable. On this basis, 374 patients were non-randomly allocated to treatment (67.65% of these individuals were prescribed escitalopram). Any participants who were antidepressant-free (or on low doses of other antidepressants) at the first assessment were started on the study medication immediately. A wash out period of two weeks was required for participants taking fluoxetine or monoamine oxidase inhibitors.

Drugs were prescribed using a flexible-dosage protocol (50-150mg/day of nortriptyline or 10-30mg/day of escitalopram). This, together with the option for non-random drug allocation, the open-label design of the study and the absence of a placebo arm were included as pragmatic study design features to make the project more inclusive and open to a wider proportion of patients. Self-reported pill counts were obtained and 98.4% of the sample reported treatment adherence.

2.3.1 Comedications

Patients were prohibited from using other psychotropic medications during the trial, excepting the occasional use of hypnotics. Any other medications taken throughout the study were recorded.

2.4 Clinical measures

2.4.1 Treatment response

2.4.1.1 Selection of appropriate measurement tool for treatment response

Antidepressant treatment response is measured as an improvement (i.e. decrease) in depressive symptoms. However, there are a number of different scales that can be used to measure depressive symptoms; no gold standard has been identified. It is unclear whether clinician-rated inventories offer a more objective measure of symptomatology (Prusoff *et al*, 1972) or lack sensitivity (Greenberg *et al*, 1992). Additionally, the appropriate weighting for each symptom encompassed within a diagnosis of depression has not been fully

established (Bagby *et al*, 2004), particularly in light of the high levels of heterogeneity between depressed patients in terms of symptomatology.

Within this context, three measurement scales were used in GENDEP; the self-report Beck Depression Inventory (BDI; Beck *et al*, 1961), the clinician-rated Hamilton Depression Rating Scale (HDRS-17; Hamilton, 1960, 1967) and the clinician-rated Montgomery-Asberg Depression Rating Scale (MADRS; Montgomery and Asberg, 1979). Each scale was administered weekly from baseline (week zero) to week twelve of treatment. At weeks zero, eight and twelve, interviews were given face-to-face by a psychiatrist and a research assistant (both trained in administration of the scales). All remaining assessments were done either by telephone or with face-to-face interviews by either a trained psychologist or psychiatrist.

Previous work exploring the three scales within the GENDEP data set concluded that MADRS outperforms the other two scales for both inter-rater reliability and accuracy of detecting depression symptoms over a range of symptom severity (Uher *et al*, 2008). Thus the MADRS has been used as the primary outcome measure throughout this thesis. The MADRS is included in Appendix A.

2.4.1.2 Selection of appropriate model for treatment response

Depression symptomatology is a continuous trait, and whilst various cut-off points to define response and remission in depression have been defined, there is a lack of consensus on which cut-off point is most appropriate (Frank *et al*, 1991; Zimmerman *et al*, 2004). Additionally, dichotomising a continuous trait results in a substantial loss of statistical power (Streiner, 2002), which is a particular concern given the evidence indicating genetic and environmental influences on antidepressant treatment response are likely to have small effect sizes. Therefore, in this thesis, antidepressant treatment response is always considered as a quantitative trait.

Given the availability of weekly data on treatment response, longitudinal measures of treatment response have been employed in this thesis, where possible. By using linear mixed models it was possible to include the repeated measures for each individual. Furthermore, these models are robust to missing data, giving unbiased estimates without the need for imputation (Lane, 2008; Mallinckrodt *et al*, 2001).

However, the inclusion of longitudinal data for each participant was not feasible within the context of analyses conducted on a genomic scale (Chapters 6 and 7). In these cases, an alternative outcome of percentage change in MADRS from baseline to week 12 of the study was used (previously described by Uher *et al*, 2010). Missing data at week 12 was imputed using linear mixed effects regression models, with fixed linear and quadratic effects of time and random effects of recruitment centre and individual. Given that treatment response was observed to be associated with age and recruitment centre, the final measure of percentage change in MADRS score was adjusted for these factors. It has been observed that within GENDEP, percentage change in MADRS score is preferable to absolute change, given its high correlation with end-of-treatment score (r=0.84), low correlation with baseline severity (r=-0.06) and aligns with clinical impressions of improvement (Uher *et al*, 2010).

2.4.2 Antidepressant side effects

2.4.2.1 Selection of appropriate measurement tool for side effects

Two inventories were used to assess antidepressant side effects in GENDEP. The self-report Antidepressant Side Effect Checklist (ASEC, Uher *et al*, 2009a) was developed as part of the GENDEP project. The checklist was designed specifically to identify side effects that have been previously identified as associated with antidepressants and indexes 21 items. The ASEC was administered weekly from week zero to week twelve. In addition, the interviewer-rated UKU (Udvalg for Kliniske Undersoegelser Side Effects Rating Scale, Lingjaerde *et al*, 1987) was administered by a psychiatrist at week zero, eight and twelve. This semi-structured interview assesses common side effects associated with psychotropic medications, indexing 48 items (14 of which were also contained within the ASEC).

The ASEC has previously been shown to correlate well with the UKU (Uher *et al*, 2009a), and given that the ASEC data available was more comprehensive, this measure was used to index adverse drug reactions throughout this thesis. The ASEC is included in Appendix B.

2.4.2.2 Selection of appropriate model for side effects

Each of the 21 items included in the ASEC were rated on a four-point scale (0 absent; 1 mild; 2 moderate; 3 severe). As moderate and severe ratings were rare in this sample, these values were collapsed to give dichotomous presence/absence outcomes per week. In addition to considering each specific side effect, a measure of total side effect burden was also constructed.

As described above for response measures, linear or logistic mixed models were used to capture the longitudinal measures of side effects, making maximum use of the weekly data available.

2.4.3 Study drop out

Study end week was recorded for each participant. The proportion of patients exiting the study each week is shown in Figure 2-1; 64.05% of participants remained in the study for the full twelve weeks.

2.5 Serum measurements

To establish circulating concentration of antidepressant and primary metabolite, serum measurements were taken using blood samples taken at week eight of the study, when steady state serum concentrations have been achieved (Hiemke & Hartter, 2000; Linder & Keck, 1998). Blood samples were taken in the morning, and patients were asked to postpone taking their morning dose of antidepressant until after the samples were collected. Serum levels of antidepressant were available for 319 patients taking escitalopram and 213 patients taking nortriptyline.

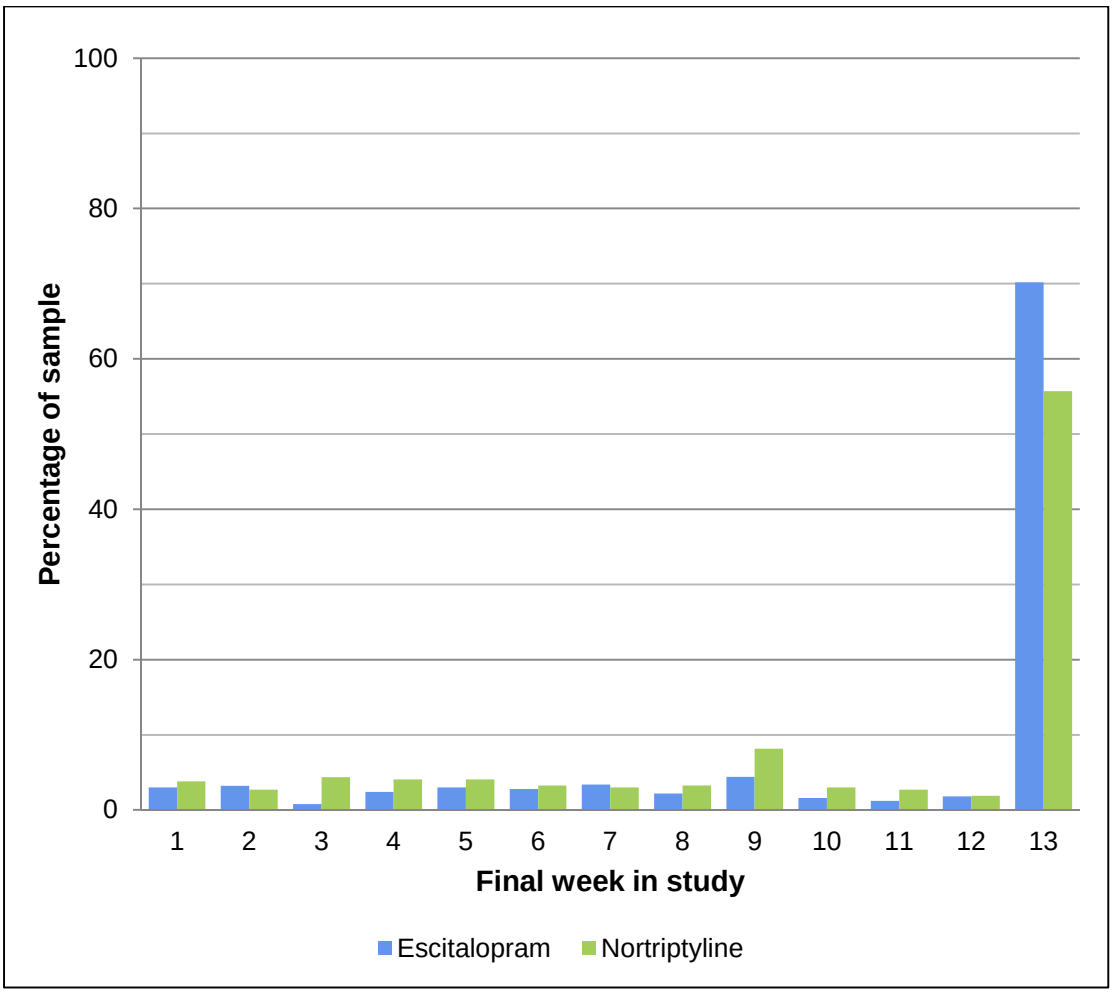

Figure 2-1: Proportion of sample exiting GENDEP study, per week per drug

All serum analyses were performed at the Department of Clinical Biochemistry, Kings College Hospital, London (UK). Escitalopram, desmethylcitalopram, nortriptyline and total 10-hydroxynortriptyline were measured using achiral turbulent flow liquid chromatography (Couchman, 2012) using an Aria Transcend TLX-II system (ThermoFisher Scientific, San Jose, USA). Detection was by tandem mass spectrometry (MS/MS) (TSQ Vantage, ThermoFisher Scientific, Hemel Hampstead, UK) in positive ionisation mode using atmospheric pressure chemical ionisation (APCI), with two selected reaction monitoring (SRM) transitions used for each analyte. The *cis*- and *trans*-isomers of 10-hydroxynortriptyline were resolved and assay calibration was based on the *cis*-isomer. Sample preparation in both cases was by protein precipitation. Samples, calibration standards, and internal quality control solutions (50 µL) were vortexed (5 min) with a protein crash solution including internal standard (250 µL), at 4 ^{o}C, in an eppendorf tube. After centrifugation

(5 min, 13,000 rpm), the supernatant was transferred to the well of a 96-well plate and 40 µL injected. The method was fully validated according to FDA guidelines (FDA, 2001).

In all cases, detection was by positive mode electrospray ionisation. Selected reaction monitoring was used (two *m/z* transitions per analyte). Each assay was calibrated using 7 calibration solutions over the following ranges; escitalopram, desmethylcitalopram, and nortriptyline: 10-500 µg/L, cis-10-hydroxynortriptyline: 10-1000 µg/L. Internal standards were: LU-10-2020, an escitalopram analogue (for escitalopram and desmethylcitalopram) and nortriptyline-D_3 (for nortriptyline and cis-10-hydroxynortriptyline). The response for trans-10-hydroxynortriptyline was assumed to be the same as that of cis-10-hydroxynortriptyline. Ion suppression or ion enhancement were not observed using standard testing procedures.

2.6 Genotypic measures

Blood sample for genetic analysis were taken at baseline. DNA was extracted from these samples and buffered in ethylenediaminetetraacetic acid (Freeman *et al*, 2003).

2.6.1 CYP450 enzymes genotyping

For CYP450 genotyping, blood samples were available for 846 participants. Genotyping was performed using the Roche AmpliChip CYP450 (Roche Molecular Diagnostics, Alameda, CA, USA), a micro-array that measures 33 variants in *CYP2D6* and two variants in *CYP2C19*. In addition, the common *17 allele observed within *CYP2C19* (Sim *et al*, 2006) was also genotyped, using a TaqMan SNP genotyping assay on the 7900HT sequence detection system (Applied Biosystems, CA, USA). Genotypes were determined using SDS software (Applied Biosystems). The nomenclature of alleles within the CYP450 enzymes is defined as by the Human Cytochrome P450 (CYP) Allele Nomenclature Database (http://www.cypalleles.ki.se/).

2.6.2 Genome-wide genotyping

For genome-wide genotyping, a total of 795 samples were available in sufficient quantity and quality. These were sent to the Centre National de Genotypage (Evry Cedex, France) and genotyped using the Illumina Human610-quad bead chip (Illumina, Inc., San Diego). This chip assays more than 610,000 single nucleotide polymorphisms (SNPs) and copy number variant markers selected to provide a comprehensive coverage across populations.

2.6.2.1 Quality control for GWAS data

Using PLINK (Purcell *et al*, 2007), standard quality control procedures were performed on the genome-wide data. Firstly, SNPs were removed if the minor allele frequency (MAF) was <0.01 or if SNP completeness was <99%. However, given GENDEP is a case-only sample, Hardy Weinberg Equilibrium was not used as a quality control filter (Wittke-Thompson *et al*, 2005). Secondly, individuals were removed if genotypic sex did not align with phenotypic sex data, or if they were identified as outliers on autosomal heterozygozity (potential indicators of sample contamination). Thirdly, using a linkage disequilibrium (LD)-pruned dataset (containing 39,658 SNPs in low LD), estimation of identity by descent was performed in PLINK to identify related individuals; in first- or second-degree relatives, the individual with less complete data was removed from the sample. Finally, individuals with genotypic completeness <95% were removed.

2.6.2.2 Population stratification

In order to account for the genetic variation that is seen across European populations (Seldin *et al*, 2006; Tian *et al*, 2008), principal component analysis was applied to the LD-pruned dataset described above, using EIGENSTRAT (Price *et al*, 2006). The first five principal components identified reach significance ($p<0.05$). The first principal component corresponded to north-south geographic locations, whilst the second corresponded to east-west locations. The third and fourth principal components jointly distinguished UK samples from the remainder of the sample. The fifth principal component showed no relationship with recruitment centre, and so the first four principal components were used as covariates to capture population

stratification. Five individuals identified as outliers were excluded (analysis with HapMap indicated Asian or African admixture).

2.6.2.3 Imputation

The Illumina Human 610-quad bead chip (Illumina, Inc., San Diego) surveys over 610,000 SNPs across the genome. However, it is possible to use imputation to estimate the values for unobserved genotypes within the study data, using data from references panels, such as the 1,000 Genomes Project (Abecasis *et al*, 2010). Patterns of linkage disequilibrium within the reference panel data can be used to extrapolate likely genotypic values within the study data (Howie *et al*, 2009). This additional imputed data increases power to detect genetic associations; this has been demonstrated for genome-wide association studies of gene expression traits (Liang *et al*, 2013). In order to impute the genotypic data within this sample, IMPUTE 2 software was used (Howie *et al*, 2011). 1000 Genomes Phase 1 integrated haplotypes (NCBI build 37) were downloaded from the IMPUTE2 website as the reference panel for the imputation (https://mathgen.stats.ox.ac.uk/impute/impute_v2.html). The study panel was the quality-controlled genotypes from the Illumina Human 610-quad bead chip. Genomic positions of variants within the study panel were converted from NCBI build 36 to build 37 coordinates with UCSC LiftOver (http://genome.ucsc.edu/cgi-bin/hgLiftOver). Any ambiguous A/T or C/G SNPs were removed, to enable automatic strand alignment options within IMPUTE2. After imputation in IMPUTE2, any genotypes with INFO scores <0.5 or MAF<0.1 were removed. The tool fcGENE was used to convert genotype probabilities into allelic dosages (http://www.imise.uni-leipzig.de/en/Groups/GenStat/Tools/index.jsp). The total number of imputed SNPs after quality control was 8,317,505 SNPs.

2.7 Transcriptomic measures

Blood samples for RNA measurement were taken at two time-points during the study; week zero and week eight. Data was available from a total of 227 participants at week zero; of these 136 participants also had data available from samples taken at week eight. The samples used for the transcriptomic analysis were taken at the same point as those used for the serum measurements (outlined in 2.5); all samples were taken

in the morning, prior to taking the morning dose of antidepressant. Typically, samples were taken at around 10am.

2.7.1 Whole blood RNA extraction and microarray expression measurement

Ten mL of blood was collected in PAXgene tubes (PreAnalytiX, Switzerland) and stored at -80 degrees. Prior to RNA extraction, PAXgene tubes were allowed to thaw for 12 hours at room temperature. RNA extraction was subsequently performed using the Qiagen PAXgene Blood miRNA Kit (PreAnalytiX) following the standard manufacturer's protocol. The purity and quantity of RNA was measured using the Nanodrop, ND1000 (Thermoscientific, Wilmington, DE), and RNA integrity numbers (RINs) were furthermore assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies, Berkshire, UK). RNA samples were then sent to the University of California Los Angeles, where genome-wide RNA expression profiling was performed using HumanHT-12 v4 Expression BeadChip microarrays (Illumina, Inc., San Diego). This microarray contains over 47,000 gene-specific probes measuring gene expression levels. Raw data was extracted using BeadStudio©. Samples were sent in four batches; assignment to batch was random.

2.7.2 Transcriptomic quality control

Two datasets were created; the first contained all week zero samples (227 samples) and the second contained samples from all individuals where data was available at both week zero and week eight (272 samples from 136 participants). These datasets were considered separately, but the same quality control metrics were used in each case. Quality control and analysis was undertaken in R (version 3.0.2). All gene expression values were log transformed. Samples with sex-incongruent expression of the *XIST* gene were removed, as were samples where inter-array correlations were more than 2 standard deviations from the mean. Detection p-values were used for probe filtering ($p<0.05$ in at least one sample), and probes displaying little variation were also removed (where the standard deviation was within the lowest quartile).

Figure 2-2 shows the quality controls details for each of the two datasets.

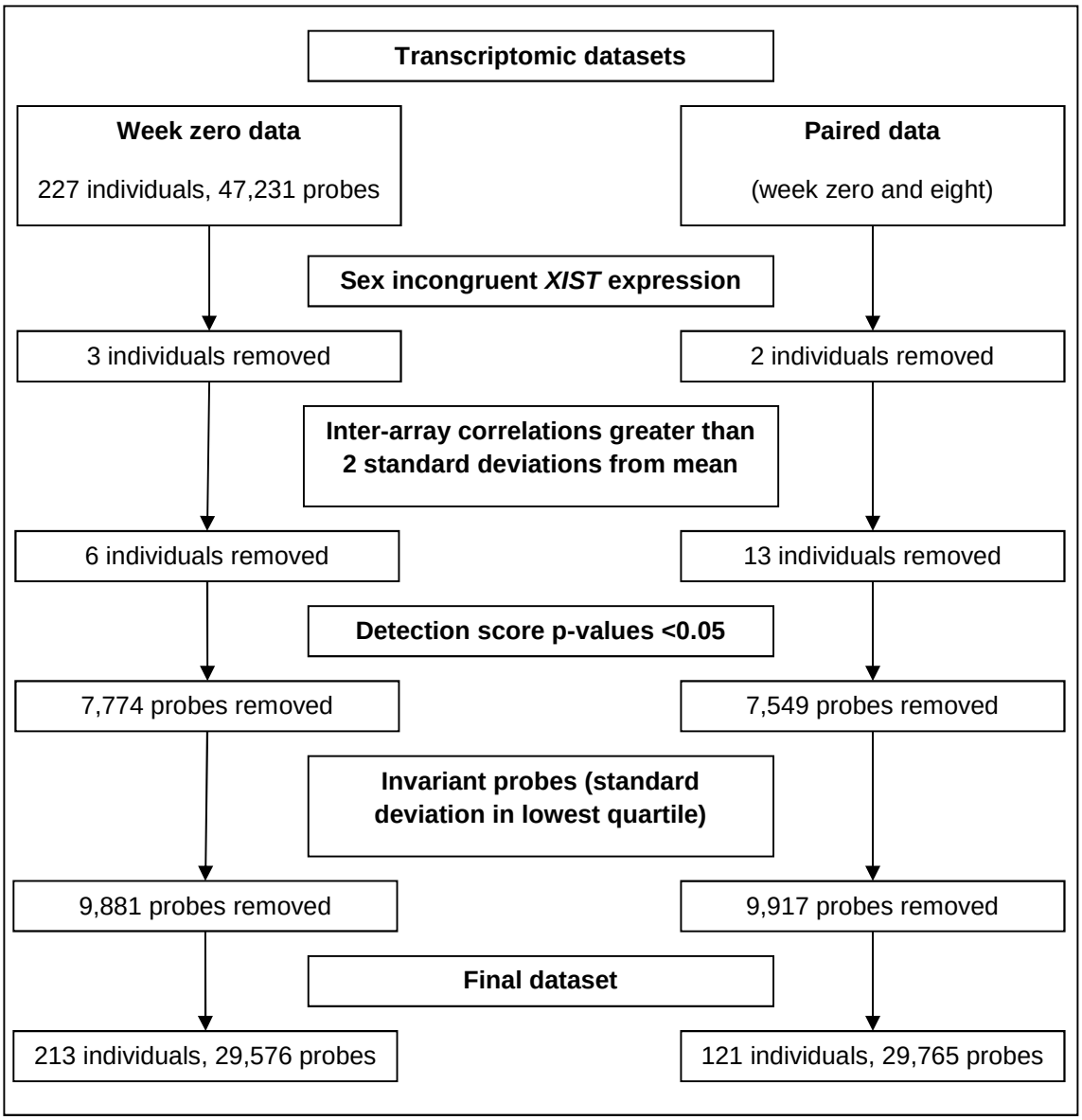

Figure 2-2: Quality control stages for transcriptomic data

After quality control, data was normalised using quantile normalisation, and ComBat (Johnson et al, 2007) was used to control for batch effects. As measurements of cell type proportions were unable for these data, deconvolution methods were used. The R package CellMix (Gaujoux and Seoighe, 2013) was used to estimate proportions of lymphocytes, neutrophils and monocytes per sample, employing previously identified cell-specific markers (Abbas *et al*, 2009).

Chapter 3 Genetic predictors of antidepressant side effects: A grouped candidate gene approach in the GENDEP study

The work presented here is published:

Hodgson K, Uher R, Crawford AA, Lewis G, O'Donovan MC, Keers R*, et al* (2014). Genetic predictors of antidepressant side effects: a grouped candidate gene approach in the Genome-Based Therapeutic Drugs for Depression (GENDEP) study. *J Psychopharmacol* **28**(2): 142-150.

3.1 Introduction

Patients taking antidepressants report a range of different medication-linked side effects. Estimates vary, but the frequency of these adverse drug reactions (ADRs) has been observed to be as high as 75% (Meijer *et al*, 2002). Given the link between ADRs and treatment non-adherence and discontinuation (Bull *et al*, 2002; Mitchell, 2006), it is important to understand the causes of these side effects. An individual's risk of ADRs is thought to be partially determined by genetic factors; identifying the genes that predict particular ADRs would enable clinicians to determine the side effect risk profile of each patient and tailor prescriptions accordingly. Despite the clinical relevance of this approach, given the number of different ADRs that are reported with antidepressants, the literature is this area is diffuse, with poor coverage of a large number of outcomes.

As noted in the Introduction, there are differences between antidepressants in terms of commonly reported ADRs, which are echoed by differences in drug receptor affinities. The majority of common antidepressant side effects are associated with one of four receptor classes; adrenergic, cholinergic, histaminergic and serotonergic. Orthostatic hypotension and dizziness are associated with antagonism of adrenergic receptors (Carruthers, 1994), whilst dry mouth, blurred vision, constipation and problems with urination result from antagonism of cholinergic receptors. Drowsiness, increased appetite and weight gain are linked to histaminergic antagonism (Lecklin *et al*, 1998; Monti *et al*, 1990) and insomnia, reduced appetite, sexual dysfunction, nausea and diarrhoea are associated with serotonergic receptors (Kennedy *et al*, 2000; Nutt, 2002; San and Arranz, 2008; Stahl, 1998). These proposed groupings are supported by the observed patterns seen between the side effect profiles and receptor affinities of various antidepressants (Hamon and Bourgoin, 2006). Tricyclic antidepressants act at cholinergic, histaminergic, serotonergic and adrenergic receptors, whilst selective serotonin reuptake inhibitors (SSRIs) target the serotonin transporter more specifically (Hamon *et al*, 2006). Consequently, one novel approach to identifying genetic predictors of ADRs is to group them based on what is known about the relationship between side effects and drug action at specific receptors.

Using these four receptor-defined categories of ADRs, this chapter aimed to investigate whether the likelihood of experiencing a particular group of side effects was linked to genetic variation in the relevant receptors.

3.2 Methods

3.2.1 Participants

The details of the GENDEP sample are described in Chapter 2. Those patients from the study with genetic information, as well as data on the experiences of ADRs (both prior to medication and for at least one week post-baseline) were selected for inclusion. Additionally, only randomly allocated patients were included in this analysis.

This is because the GENDEP study design involved partial-random allocation to antidepressant medication; individuals with previous contra-indications were non-randomly allocated to the alternative antidepressant (36% of the sample). Non-random allocation to treatment is based on the assumption that patients are likely to experience the side effect again if given the same medication, and less likely to experience the side effect when taking the alternative medication. If these assumptions are true, the ability to detect genetic associations with ADRs is reduced in non-randomly allocated patients.

Figure 3-1 shows the patients from GENDEP that were included in the analyses presented in this chapter.

As described in Chapter 2.3, two antidepressants with divergent actions were used in GENDEP: escitalopram and nortriptyline. Briefly, escitalopram is an SSRI with a very high affinity for the serotonin transporter and negligible affinity for other receptors in the brain (Sanchez *et al*, 2003). In contrast, nortriptyline is a tricyclic antidepressant, with preferential action at the noradrenergic transporter, but also shows serotonergic, adrenergic, muscarinic and histaminergic action (Sanchez and Hyttel, 1999).

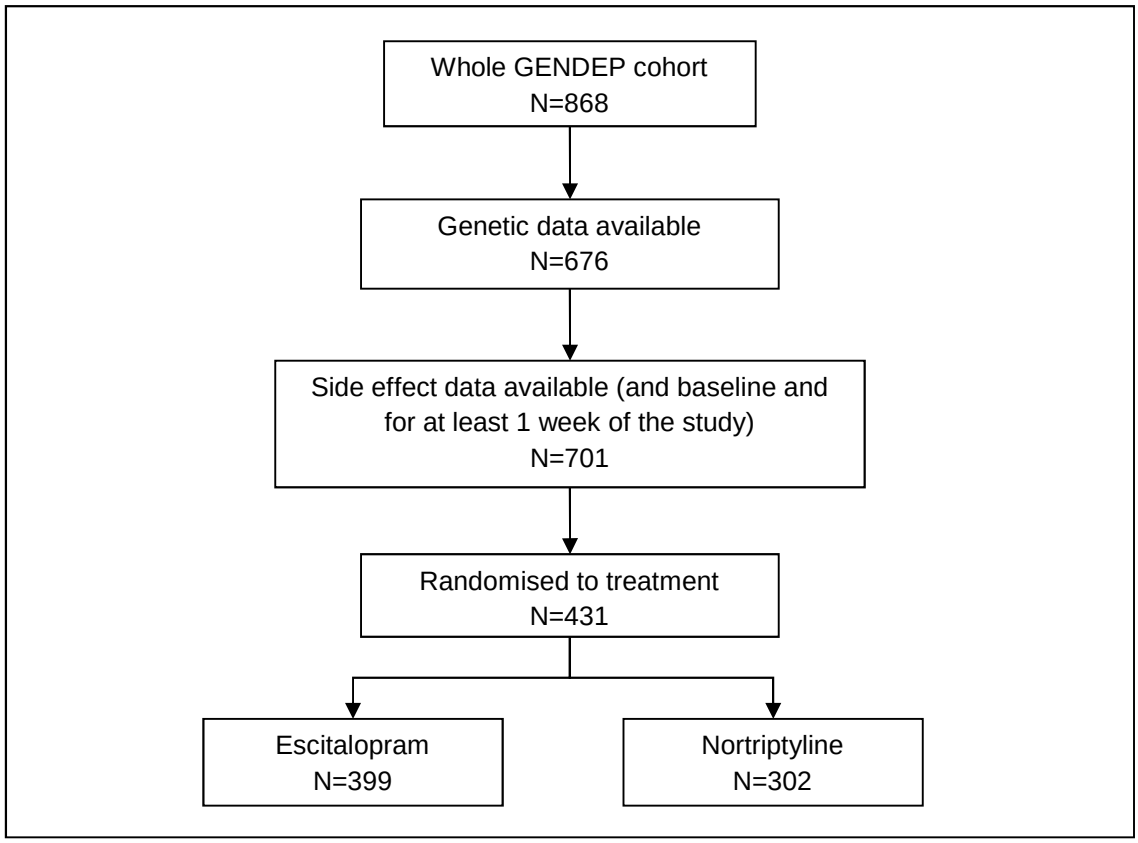

Figure 3-1: Sample included in Chapter 3

3.2.2 Measures

Side effects were measured weekly, using the self-report Antidepressant Side Effect Checklist (ASEC), as described in Chapter 2.4.2. The ASEC measures 21 side effects. From these, four categories were derived, based on the putative pharmacological basis of the side effects; serotonergic, adrenergic, histaminergic or cholinergic, as shown in Table 3-1. Adverse reactions where the pharmacological basis is unclear were excluded, as there may be several interacting processes involved. These included palpitations, headache, sweating, increased temperature, tremor, disorientation, yawning.

Although each item on the ASEC is rated on a four-point scale (0 absent; 1 mild; 2 moderate; 3 severe), moderate and severe ratings were uncommon (Uher *et al*, 2009a). Thus, for each outcome group (adrenergic, cholinergic, histaminergic or serotonergic) and for each week of the study where ratings were

available, ADRs were coded as present if the patient scored any of the side effects in the grouping at one or above.

Table 3-1: ASEC items grouped according to physiological pathway

Serotonergic	Adrenergic	Cholinergic	Histaminergic
Nausea or vomiting	Feeling light-headed on standing	Problems with urination	Drowsiness
Diarrhoea	Feeling like the room is spinning	Blurred vision	Increased appetite
Decreased appetite		Constipation	Weight gain
Insomnia		Dry mouth	
Sexual problems			

3.2.3 Genotyping

Using the quality-controlled genome-wide data described in Chapter 2.6.2, genotypic information was extracted for twenty-four candidate genes. The candidate genes were selected across the four receptor systems proposed to underlie each side effect group. For each gene, 1000bp up and downstream of the coding sequence was included, to capture proximal promoter sequences (Solovyev and Shahmuradov, 2003). Only variants with a minor allele frequency greater than 5% were included.

Additional variants within genes of particular interest in depression were also genotyped in the sample (as described in Uher *et al*, 2009b). Four of these genes had been selected as candidates in this study and this additional genotypic information supplemented the genome-wide data. In the serotonin transporter gene (*SLC6A4*), the serotonin-transporter-gene-linked polymorphic region (*5-HTTLPR*), and the putative functional marker rs25531 that lies within this region were both genotyped in a two-stage method. Two further polymorphic microsatellite markers (STin2 and STin4) within the *SLC6A4* gene were also genotyped, and after examination of the distribution of repeats, both STin2 and STin4 were categorised as either “long” (STin2: 12 repeats, STin4: 8-9 repeats) or “short” (STin2: 9-10 repeats, STin4: 5-7 repeats). All other additional SNPs were genotyped using the SNPlex method (Applied Biosystems, Inc., Foster City, CA) with genotype discrimination achieved using the Applied Biosystems ABI 3130 sequencer. Data was exported

and analysed with GENEMAPPER software (Applied Biosystems, version 4.0), see Huezo-Diaz *et al* (2009) for further details and quality control thresholds applied.

In total, genotypic data was available for 436 markers from twenty candidate genes (see Table 3-2 for further details).

Table 3-2: Candidate genes selected for ADR analysis in GENDEP sample. (Meff: effective number of comparisons, as calculated using SNPSpD)

Candidate gene	Associated side effects	No. of markers available	Gene-wide significance threshold (0.05/Meff)
SLC6A2	Adrenergic	47	0.00117
ADRA1A	Adrenergic	53	0.00103
ADRA1B	Adrenergic	14	0.00415
ADRA1D	Adrenergic	7	0.00810
SCL6A4	Serotonergic	21	0.00320
HTR1A	Serotonergic	3	0.04808
HTR1B	Serotonergic	0	
HTR1D	Serotonergic	0	
HTR1E	Serotonergic	9	0.00668
HTR1F	Serotonergic	0	
HTR2A	Serotonergic	51	0.00110
HTR2B	Serotonergic	2	0.02666
HTR2C	Serotonergic	14	0.00520
HTR3A	Serotonergic	9	0.00922
HTR3B	Serotonergic	5	0.01228
HTR3C	Serotonergic	2	0.03182
HTR3D	Serotonergic	5	0.01313
HTR3E	Serotonergic	2	0.02582
CHRM1	Cholinergic	3	0.01730
CHMR2	Cholinergic	32	0.00191
CHRM3	Cholinergic	119	0.00044
CHRM4	Cholinergic	0	
CHRM5	Cholinergic	16	0.00449
HRH1	Histaminergic	22	0.00257

3.2.4 Statistical Analysis

3.2.4.1 Regression model

For each of the four groupings (serotonergic, adrenergic, histaminergic and cholinergic), logistic regression was used to test the effect of genotype on the presence or absence of side effects under an additive genetic model. For genes located on the X chromosome, males were treated as homozygotes. Data on ADRs from all 12 weeks of the study were included in the analyses; therefore a Huber-White sandwich estimator of variance was used to account for the use of repeated measurements (Kuzman *et al*, 2008). This estimator relaxes the assumption of independent observations by providing standard errors that are robust to intra-individual correlations (Kent, 1982).

To examine only those side effects that emerged as a result of antidepressant treatment, baseline ADR ratings (prior to treatment initiation) were controlled for. Additionally, given the effect of depression severity on ADR endorsement (Uher *et al*, 2009a), current and baseline ratings on the Montgomery-Asberg Depression Rating Scale (Montgomery *et al*, 1979) were entered as covariates in all analyses. Other covariates included age, sex, study week, dose (standardised to allow comparison between drugs), and drug (when applicable).

3.2.5 Subgroup analyses

Escitalopram and nortriptyline have different receptor affinities, and it was hypothesized that the association between genotype and side effect may be specific to patients taking medications with direct action at the relevant receptors. Thus, for each of the four ADR outcomes, associations between genotype and ADR ratings were examined for the following groups: A) all patients randomly allocated to their medication; B) patients randomly allocated to escitalopram; C) patients randomly allocated to nortriptyline. All statistical analyses were performed in STATA version 10.1 (STATACorp LP., College Station, TX)

3.2.6 Correction for multiple testing

In order to control for multiple tests, the effective number of comparisons was calculated based on the procedure described by Li and Ji (2005), using SNPSpD software (Nyholt, 2004). This takes into account the linkage disequilibrium that exists between SNPs. For each ADR outcome, 436 SNPs were tested for association, equivalent to 427.13 effective comparisons. This effective number of comparisons (Meff) was used to calculate an outcome-wide significance level (0.05/Meff) of p=1.17 x 10^{-4}.

In order to identify any SNPs that may not reach outcome-wide significance, but do reach significance on a gene-wide scale, the effective number of comparisons per gene was also calculated (as shown in Table 3-2).

To estimate the posterior probability of true positive findings in the context of multiple non-independent tests, false discovery rate q-values were calculated using the Benjamini and Hochberg step-up procedure (Theisen *et al*, 2004) and QVALUE software (Risselada *et al*, 2010). These q-values indicate the proportion of false positives amongst those tests that reach significance.

3.2.7 Replication in GenPod

The GenPod study was used as a replication sample to follow up significant associations that were observed in the GENDEP sample. GenPod is a UK-based multi-centre trial where depressed patients are randomised to receive either citalopram (of which escitalopram is the (S)-stereo-isomer) or reboxetine (a specific noradrenergic reuptake inhibitor, which unlike nortriptyline, has a weak affinity for muscarinic, H1 histaminergic or adrenergic α-1 receptors (Boothman *et al*, 2006). Patients were followed over a period of 12 weeks, with assessments taken at baseline, week 6 and 12 of treatment. Side effects were measured using a modified version of the Toronto Side Effects Scale (TSES; Lesch *et al*, 1994), included in Appendix C, and depression severity was measured with the self-report Beck Depression Inventory (BDI). Samples were genotyped at the University of Geneva Medical School (Geneva, Switzerland) using the Illumina Human660W-Quad BeadChip (Illumina Inc., San Diego, CA). Further details are published elsewhere

(Cremers *et al*, 2004). From a sample of 601 patients, 474 had both genetic and ADR data available (50.2% of these were taking citalopram).

3.3 Results

3.3.1 Adverse drug reaction profiles in GENDEP

Figure 3-2 demonstrates the pattern of side effects in GENDEP across the 12-week trial, split by medication taken. The expected relationships are observed, with adrenergic, histaminergic and cholinergic outcomes more frequent amongst patients taking nortriptyline, and serotonergic adverse effects more common amongst those on escitalopram. However, this difference is relative; all groups of side effects are reported with both antidepressants. Further, there are notably high rates of side effect endorsement at baseline (study week zero), when all patients were unmedicated.

3.3.2 Genetic associations in GENDEP

All associations reaching gene-wide significance are shown in Table 3-3 and Table 3-4.

3.3.2.1 Cholinergic adverse reactions

No SNPs reached outcome-wide significance when considering cholinergic adverse reactions. A gene-wide significant association was observed with a SNP in *HTR2B* in the whole sample, whilst for individuals taking escitalopram, there was a gene-wide significant association with a SNP in cholinergic receptor *CHRM3*. In patients taking nortriptyline, there were associations in both *CHRM3* and *HTR2B* at gene-wide levels, however at different variants to those identified in the escitalopram-specific (*CHRM3*) or randomised (*HTR2B*) analyses. The linkage disequilibrium between the implicated SNPs in each of these genes was low.

3.3.2.2 Adrenergic adverse reactions

No outcome-wide significant associations were observed with adrenergic side effects. In the whole sample analysis, one SNP in *HTR3E* reached gene-wide significance. This same SNP (rs7432211) was also associated at a gene-wide level in the nortriptyline-specific analysis. While escitalopram has no adrenergic action, three SNPs in three different genes were associated at a gene-wide threshold (*CHRM3*, *ADRA1A* and *HTR3C*) in the escitalopram-specific analysis.

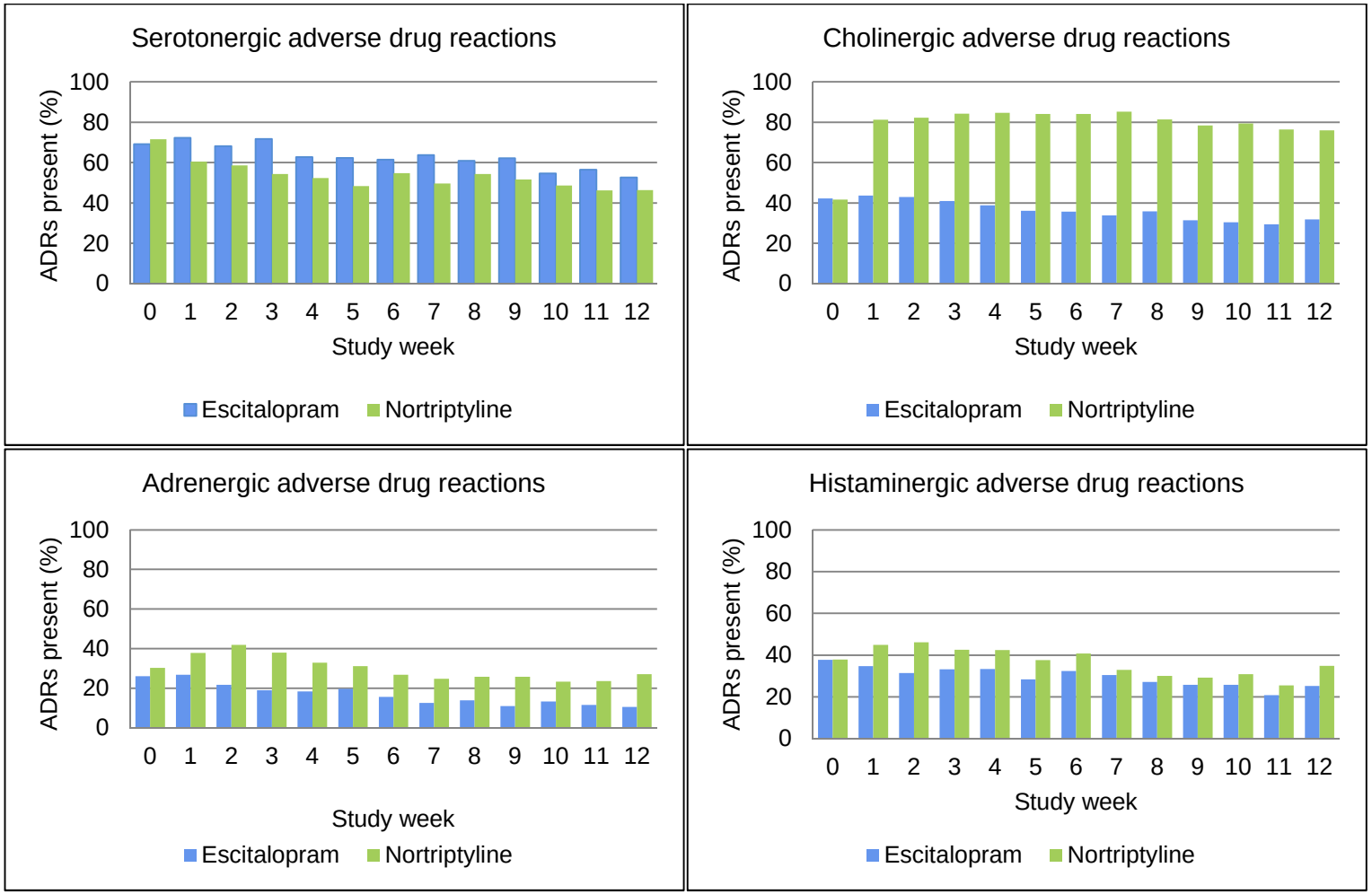

Figure 3-2: Adverse drug reactions per week in GENDEP, separated by drug.

Table 3-3: Genetic associations with cholinergic, adrenergic and histaminergic adverse drug reactions in GENDEP sample.

Analysis: A=whole sample; B=patients taking escitalopram; C=patients taking nortriptyline. SNPs that are present in more than one analysis are marked in superscript (1=also in analysis A, 2=also in analysis B, 3=also in analysis C). MAF= minor allele frequency, n=number of individuals, Obs=total number of observations across all 12 weeks. Odds Ratios are per minor allele.

Analysis	Gene	Gene-wide sign	C'some	SNP	Location	Allele	MAF	n	Obs	p	q	Odds Ratio	95% Confidence Interval	
Cholinergic														
A	HTR2B	0.02666	2	rs10194776	231980019	T/C	0.40	362	3290	0.0245	0.4357	1.37	1.04	1.81
B	CHRM3	0.00044	1	rs1431718	239879553	T/C	0.41	184	1764	2.92E-04	0.0479	1.87	1.33	2.62
C	CHRM3	0.00044	1	rs11578320	239906616	C/T	0.07	178	1526	1.95E-04	0.0826	5.41	2.23	13.17
	HTR2B	0.02666	2	rs4973377	231981992	A/G	0.17	178	1526	0.0091	0.6502	0.43	0.23	0.81
Adrenergic														
A	HTR3E	0.02582	6	rs7432211[(3)]	183819155	C/T	0.41	361	3276	0.0200	0.4634	1.38	1.05	1.81
B	CHRM3	0.00044	1	rs685548	239994906	T/G	0.39	183	1752	3.38E-04	0.1006	0.48	0.32	0.72
	ADRA1D	0.00810	20	rs6084670	4222509	C/A	0.24	183	1752	0.0044	0.3115	0.48	0.29	0.80
	HTR3C	0.03182	3	rs6808122	183772821	G/A	0.37	183	1752	0.0093	0.3892	0.58	0.39	0.88
C	HTR3E	0.02582	3	rs7432211 [(1)]	183819155	C/T	0.41	178	1524	0.0032	0.3920	1.72	1.20	2.46
Histaminergic														
A	HTR3D	0.01313	3	rs6792482	183754029	C/T	0.44	362	3289	0.0078	0.7847	1.37	1.09	1.73
B														
C														

Table 3-4: Genetic associations with serotonergic adverse drug reactions in GENDEP sample.

Analysis: A=whole sample; B=patients taking escitalopram; C=patients taking nortriptyline. SNPs that are present in more than one analysis are marked in superscript (1=also in analysis A, 2=also in analysis B, 3=also in analysis C). MAF=minor allele frequency, n=number of individuals, Obs=total number of observations across all 12 weeks. Odds Ratios are per minor allele.

Analysis	Gene	Gene-wide sign	C'some	SNP	Location	Allele	MAF	n	Obs	p	q	Odds Ratio	95% Confidence Interval	
						Serotonergic								
A	HTR2C	0.00520	X	**rs6644093[2]**	**114064023**	**T/G**	**0.15**	**362**	**3288**	**7.43E-05**	**0.0252**	**1.72**	**1.31**	**2.25**
				rs4911871[2]	113997140	G/A	0.21	362	3288	0.0016	0.1074	1.49	1.16	1.91
				rs12846241	113854086	G/T	0.18	362	3288	0.0021	0.1074	1.52	1.17	1.99
				rs12690355	113910850	G/A	0.18	361	3280	0.0022	0.1074	1.53	1.64	2.01
				rs2428700	114010664	A/G	0.14	361	3284	0.0031	0.1074	0.63	0.46	0.86
				rs4332303	114047867	T/C	0.14	362	3288	0.0031	0.1074	0.63	0.46	0.86
				rs5946005	114082535	G/A	0.14	362	3288	0.0031	0.1074	0.63	0.46	0.86
				rs5988087	113934856	T/C	0.16	362	3288	0.0032	0.1074	0.64	0.48	0.86
				rs11167436	113944060	A/C	0.16	362	3288	0.0032	0.1074	0.64	0.48	0.86
	CHRM2	0.00191	7	rs1364403	136588827	T/C	0.30	362	3288	0.0020	0.1074	0.65	0.49	0.85
B	HTR2C	0.00520	X	rs6644093[1]	114064023	T/G	0.15	184	1763	6.07E-04	0.2642	2.02	1.35	3.01
				rs4911871[1]	113997140	G/A	0.21	184	1763	0.0032	0.4062	1.70	1.19	2.42
	CHRM1	0.01730	11	rs2067477	62678306	T/G	0.09	184	1763	0.0030	0.4062	0.55	0.33	0.89
				rs2075748	62688269	T/C	0.21	184	1763	0.0115	0.5204	1.74	1.13	2.68
	HTR2B	0.02666	2	rs4973377	231981992	A/G	0.15	184	1763	0.0141	0.5204	1.82	1.13	2.94
C														

3.4.1.1 Histaminergic adverse reactions

In the whole sample analysis, one SNP in *HTR3D* was associated at gene-wide significance. No associations were seen, even at gene-wide significance thresholds, in the drug-specific analyses.

3.4.1.2 Serotonergic adverse reactions

In the whole sample analysis of serotonergic side effects, one SNP in *HTR2C* (rs6644093) achieved outcome-wide significance (OR=1.72, 95% CI=1.13-2.25, p=7.43 x 10^{-5}). The occurrence of serotonergic side effects for each genotype at this SNP is shown in Figure 3-3.

Given the significance of this association, the analysis was repeated, restricted to those patients where drug adherence could be confirmed using plasma concentrations (n=204, see Chapter 2.5 for further details). The association between rs6644093 and serotonergic ADRs was verified in this restricted analysis (OR=1.54, 95% CI=1.10-2.17, p=1.15 x 10^{-2}). Eight further SNPs in *HTR2C* reached gene-wide significance in the whole sample analysis, along with one variant within *CHRM2*.

In the escitalopram-specific analysis, two *HTR2C* SNPs, two *CHRM1* SNPs and one SNP in *HTR2B* reached gene-wide significance with serotonergic ADRs. Amongst patients taking nortriptyline, no SNPs were associated with serotonergic outcomes at a gene-wide level.

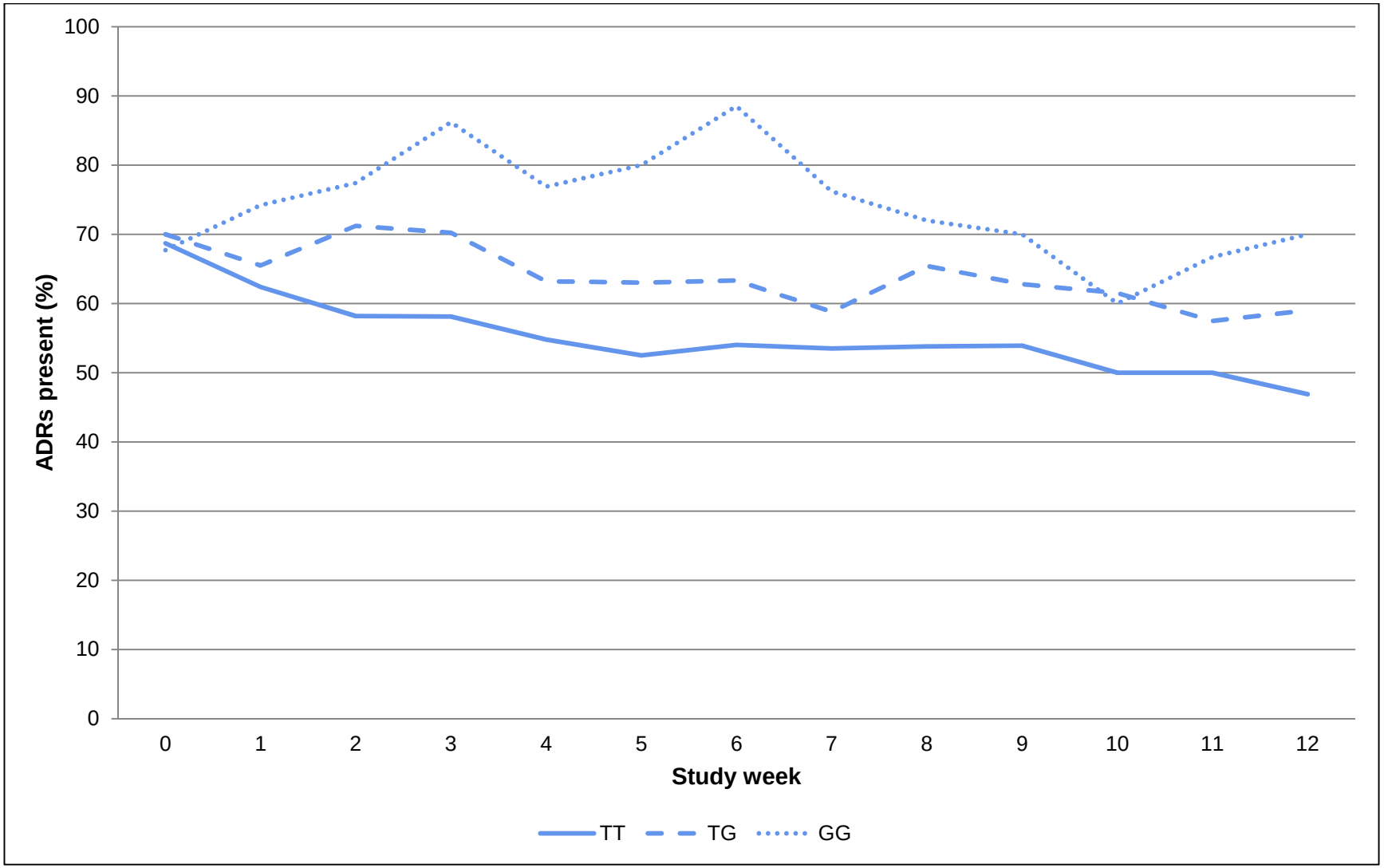

Figure 3-3: Percentage of GENDEP patients reporting serotonergic ADRs, by rs6644093 genotype

3.4.2 Replication analysis in GenPod

Serotonergic ADRs identified in GENDEP using the ASEC were matched to items included in the TSES used in GenPod (Table 3-5). However, data on sexual problems was not available for the whole sample in GenPod. In the two samples, a similar longitudinal profile of decreasing prevalence with time was observed for the serotonergic ADRs, although in GenPod, these ADRs were not more common amongst patients taking an SSRI. Furthermore serotonergic ADRs were more frequently reported in the GenPod sample than in GENDEP (see Figure 3-4).

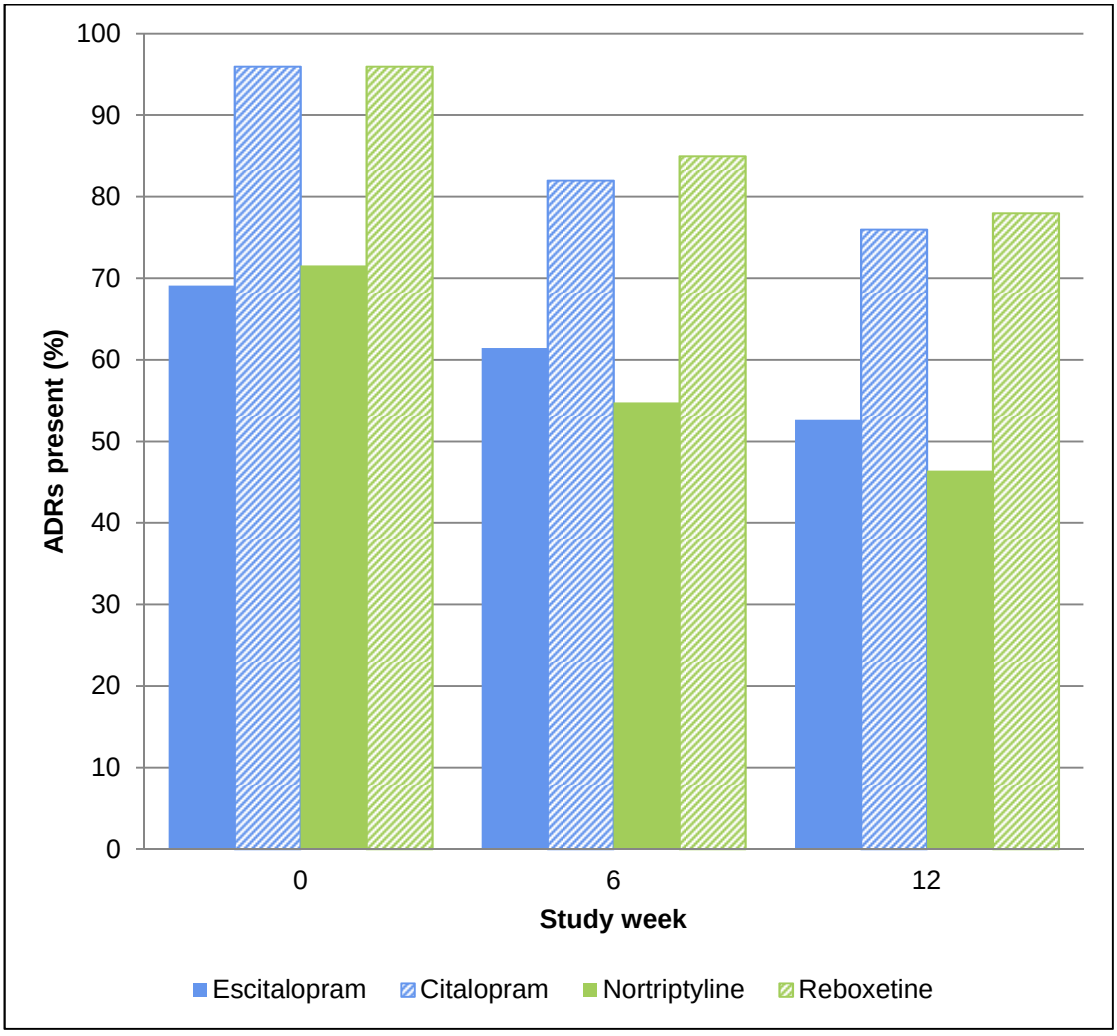

Figure 3-4: Comparison of frequency of serotonergic ADRs in GENDEP and GenPod.

Side effect scoring, genetic models and analysis framework mirrored those used in the analysis of data from GENDEP. Given that reboxetine does not display direct serotonergic activity, the analysis focussing on the subgroup of patients taking citalopram was of primary interest. In

GenPod, 13 SNPs were available within the *HTR2C* gene (8 of which were also genotyped in the GENDEP sample). For the most significant association in the GENDEP sample (rs6644093), the association in the GenPod sample was not significant (OR=0.91, 95% CI=0.55-1.52, *p*=0.723). Furthermore, no other variants in the *HTR2C* gene showed association with the occurrence of serotonergic side effects, when corrected for the effective number of SNPs tested within HTR2C (0.05/Meff; $p<6.50 \times 10^{-3}$), or when using a nominal significance ($p<0.05$). These results are shown in Table 3-6. The analyses of the whole GenPod sample and the reboxetine-specific analysis are shown in Table 3-7 and Table 3-8, again no significant associations were found.

Table 3-5: Serotonergic adverse drug reactions as indexed in GENDEP and GenPod

GENDEP (measured using the ASEC)	**GenPod** (measured using the modified TSES)
Nausea or vomiting	Felt sick or nauseous
Diarrhoea	Diarrhoea
Decreased appetite	Noticed changes in the way food tastes
Insomnia	Difficulty sleeping
Sexual Problems	*(None measured in whole sample)*

Table 3-6: Genetic association with serotonergic adverse drug reactions in GenPod patients taking citalopram.

MAF= minor allele frequency, n=number of individuals, Obs=total number of observations across 2 timepoints. Odds Ratios are per minor allele. Significance threshold =p<6.50 x 10-3, corrected for the effective number of SNPs tested within HTR2C.

Gene	C'some	SNP	Location	Allele	MAF	n	Obs	p	Odds Ratio	95% Confidence Interval	
HTR2C	X	rs5988087	113841112	A/G	0.18	237	459	0.713	0.93	0.63	1.37
		rs11167436	113850316	A/G	0.18	237	459	0.713	0.93	0.63	1.37
		rs4911871	113903396	G/A	0.21	237	459	0.347	1.29	0.76	2.2
		rs4332303	113954123	A/G	0.18	237	459	0.497	0.87	0.59	1.29
		rs6644093	113970279	A/G	0.16	237	459	0.723	0.91	0.55	1.52
		rs5946005	113988791	G/A	0.18	237	459	0.497	1.14	0.78	1.68
		rs10875535	114033435	G/A	0.05	235	455	0.735	0.86	0.35	2.08
		rs1801412	114048960	C/T	0.05	237	459	0.729	0.85	0.35	2.08
		rs1414324	114054754	A/G	0.19	237	459	0.611	0.91	0.62	1.33
		rs1335617	114084871	G/A	0.19	237	459	0.611	1.1	0.75	1.62
		rs5987834	114092371	A/G	0.19	237	459	0.611	0.91	0.62	1.33
		rs6579571	114122419	G/A	0.13	237	459	0.197	1.42	0.83	2.44
		rs638376	114139497	G/A	0.42	237	459	0.275	1.22	0.85	1.73

Table 3-7: Genetic association with serotonergic adverse drug reactions in GenPod (whole sample analysis).MAF= minor allele frequency, n=number of individuals, Obs=total number of observations across 2 timepoints. Odds Ratios are per minor allele. Significance threshold =p<6.50 x 10-3, corrected for the effective number of SNPs tested within HTR2C.

Gene	C'some	SNP	Location	Allele	MAF	n	Obs	p	Odds Ratio	95% Confidence Interval	
HTR2C	X	rs5988087	113841112	A/G	0.18	473	900	0.914	0.98	0.72	1.34
		rs11167436	113850316	A/G	0.18	473	900	0.914	0.98	0.72	1.34
		rs4911871	113903396	G/A	0.20	473	900	0.188	1.28	0.89	1.84
		rs4332303	113954123	A/G	0.17	473	900	0.709	0.94	0.69	1.29
		rs6644093	113970279	A/G	0.15	473	900	0.553	0.89	0.61	1.30
		rs5946005	113988791	G/A	0.17	473	900	0.709	1.06	0.78	1.45
		rs10875535	114033435	G/A	0.06	470	895	0.158	0.70	0.42	1.15
		rs1801412	114048960	C/T	0.06	472	898	0.149	0.69	0.42	1.14
		rs1414324	114054754	A/G	0.17	473	900	0.867	0.97	0.71	1.33
		rs1335617	114084871	G/A	0.17	473	900	0.867	1.03	0.75	1.40
		rs5987834	114092371	A/G	0.17	473	900	0.867	0.97	0.71	1.33
		rs6579571	114122419	G/A	0.13	473	900	0.656	1.09	0.74	1.61
		rs638376	114139497	G/A	0.41	473	900	0.891	1.02	0.79	1.31

Table 3-8: Genetic association with serotonergic adverse drug reactions in GenPod (reboxetine-specific analysis).

MAF= minor allele frequency, n=number of individuals, Obs=total number of observations across 2 timepoints. Odds Ratios are per minor allele. Significance threshold =p<6.50 x 10-3, corrected for the effective number of SNPs tested within HTR2C.

Gene	C'some	SNP	Location	Allele	MAF	n	Obs	p	Odds Ratio	95% Confidence Interval	
HTR2C	X	rs5988087	113841112	A/G	0.165	236	441	0.532	1.18	0.7	1.98
		rs11167436	113850316	A/G	0.165	236	441	0.532	1.18	0.7	1.98
		rs4911871	113903396	G/A	0.183	236	441	0.416	1.22	0.75	2
		rs4332303	113954123	A/G	0.156	236	441	0.599	1.16	0.67	2
		rs6644093	113970279	A/G	0.143	236	441	0.636	0.87	0.49	1.55
		rs5946005	113988791	G/A	0.156	236	441	0.599	0.86	0.5	1.49
		rs10875535	114033435	G/A	0.062	235	440	0.129	0.61	0.32	1.16
		rs1801412	114048960	C/T	0.06	234	439	0.121	0.6	0.31	1.15
		rs1414324	114054754	A/G	0.153	236	441	0.516	1.2	0.69	2.07
		rs1335617	114084871	G/A	0.153	236	441	0.516	0.83	0.48	1.44
		rs5987834	114092371	A/G	0.153	236	441	0.516	1.2	0.69	2.07
		rs6579571	114122419	G/A	0.119	236	441	0.24	0.69	0.38	1.28
		rs638376	114139497	G/A	0.398	236	441	0.248	0.8	0.55	1.17

3.5 Discussion

Identifying predictors of antidepressant side effects is an important aim with clear clinical utility, given both the prevalence of ADRs and their impact on treatment adherence. Here, a novel approach was used to classify ADRs using the wealth of evidence linking drug action at specific receptors with the emergence of particular side effects. ADRs were grouped into four receptor-defined categories, adrenergic, cholinergic, histaminergic and serotonergic, and their relationship to genetic markers within relevant receptors was studied.

No clear evidence of genetic association with adrenergic, cholinergic or histaminergic side effects were seen in the GENDEP sample, even when only considering those patients taking medication with stronger affinity for these receptors (nortriptyline). Nonetheless, there was evidence that an individual's risk of experiencing antidepressant side effects with a serotonergic basis may be influenced, in part, by *HTR2C*, with a significant relationship observed looking in the whole sample, and suggestive gene-wide associations in the escitalopram-specific analysis. However, this finding was not replicated in the GenPod sample.

3.5.1 Adrenergic, cholinergic and histaminergic outcomes

Assuming genetic factors do play a role in determining an individual's liability to these side effects, there are several possible reasons why genetic associations were not observed with adrenergic, histaminergic or cholinergic outcomes. Firstly, individual differences in the occurrence of the ADRs considered may not be determined by those genetic variants tested in this study. Whilst our knowledge of the pharmacology underlying side effects allows us to make rational choices when selecting candidate genes, this study focused only on the receptors themselves; it may be that the genes involved may be those that encode downstream effectors such as G-proteins.

Secondly, the study may have failed to demonstrate genetic associations with adrenergic, cholinergic or histaminergic side effects because the pharmacologically defined groupings used can only be a simplified model of the complex biological interactions leading to ADRs.

Finally, the study may have lacked sufficient power to detect genetic effects for adrenergic, cholinergic or histaminergic effects. If the association between genotype and ADRs is only robust when the patient is taking medication with drug action at the relevant receptor, as only nortriptyline displays adrenergic, cholinergic or histaminergic action, genetic associations might be limited to the nortriptyline-only sample for these three outcomes. Given the reduced sample sizes available when performing drug-specific analyses, these may have been underpowered.

3.5.2 *HTR2C* and serotonergic side effects

Considering serotonergic side effects, in addition to the outcome-wide significant association with a SNP within *HTR2C* (rs6644093) in the whole GENDEP sample, supportive evidence of association was also observed with further SNPs in *HTR2C* at a gene-wide level across the escitalopram-specific analysis. Escitalopram acts exclusively at the serotonin transporter, altering serotonin levels at the synapse and so indirectly impacting upon the postsynaptic 5HT-2C receptor. In contrast, nortriptyline not only displays modest affinity for the serotonin transporter, but also has an antagonistic action at the 5HT-2C receptor. It is unclear what the net effect of these two actions of nortriptyline would be, particularly given that there is likely to be an interaction between these two processes (Bostwick, 2010). Nonetheless, the association with genetic variation in *HTR2C* was seen most convincingly in the whole of the GENDEP sample, which included patients on both medications. This may reflect the increased power of this analysis due to the larger sample size.

A relationship between *HTR2C* polymorphisms and serotonergic ADRs is consistent with what is known about the physiology of the side effects included in this grouping. Not only has experimental manipulation at the 5HT-2C receptor been connected to gut motility (Fujitsuka *et al*, 2009), sleep regulation (Uhr *et al*, 2000), sexual dysfunction (Heils *et al*, 1996) and feeding

behaviour (San *et al*, 2008), but antidepressants with prominent antagonistic effects at 5HT-2C are associated with reduced levels of serotonergic side effects, relative to other antidepressants (Hu *et al*, 2006; Lesch *et al*, 1996). Indeed, in one study, nausea, insomnia and diarrhoea were reported at higher levels by those taking placebo than by patients on agomelatine (Kennedy *et al*, 2002).

However, when considering the GenPod sample, no association was found between serotonergic ADRs and variation in the *HTR2C* gene, in either the whole sample, nor in drug-specific samples. Whilst the two studies were similar in design, this failure to replicate may have been due to differences that can be seen in the rates of reporting ADRs between the two samples; in GenPod, serotonergic ADRs are not only more frequently reported, but also do not show differences in reporting frequencies between patients taking serotonergic versus noradrenergic antidepressants. Another difference between the two studies is the frequency with which ADRs were measured after baseline (in GENDEP, twelve weekly measurements were taken, whilst in GenPod measurements were only taken at week six and week twelve). However, in a supplementary analysis of the GENDEP data, using only the information collected on week six and week twelve of the study, the positive association between serotonergic ADRs and genetic variation in the *HTR2C* was still observed (see Appendix D). This indicates that the non-replication in the GenPod sample cannot be attributed to the differences in the frequency of data collection.

Other sources of variation between the studies include; differences in terms of the type of noradrenergic antidepressants administered (nortriptyline versus reboxetine) and the different scales used to index the ADRs (the ASEC versus the TSES). Nonetheless, it remains that the finding could not be replicated in GenPod and thus, may be a false positive association.

3.5.3 Limitations

The region surrounding each gene was tightly defined as 100bp up and downstream from the coding sequence, as described by Solovyev and Shahmuradov (2003). However, it may be that

more distant variants could play a role in determining individual risk to antidepressant side effects. To address this, a supplementary analysis was undertaken, where the region surrounding each candidate gene was expanded to cover 50KB up and downstream. This expanded approach was only possible within the GENDEP sample. The details of this analysis are included in Appendix D.

The association of rs6644093 with serotonergic ADRs remained significant in this expanded analysis. Only one of the additional variants showed significant association with any of the four phenotypes (serotonergic, cholinergic, adrenergic and histaminergic ADRs) tested, when performing whole sample and drug-specific analyses. This variant (rs6467694) is 7318bp downstream from the CHRM2 gene transcript (this includes the UTR), and showed significant association with cholinergic outcomes in the whole sample analysis (OR=3.09, 95% CI=1.81-5.29, p=3.75 x 10^{-5}), but no significant association was seen in the drug-specific analysis. This indicates that the cholinergic side effects observed with antidepressants are linked to genetic variation in the muscarinic cholinergic receptor 2, which aligns with the known role of cholinergic receptors in these side effects. However, no other SNPs within this gene showed evidence of association (using the gene-wide suggestive threshold of p<0.00107), and the variant did not reach gene-wide significance in either of the drug-specific analyses. This is surprising given the highly drug-specific profile of cholinergic side effects, but may reflect limitations in statistical power due to sample size in these subset analyses. As it was not possible to replicate this analysis within the GenPod dataset, and the lack of suggestive signal from other SNPs included in this gene, the association between *CHRM2* and cholinergic side effects should be treated with caution until further evidence of association is found.

A fundamental issue when addressing drug-induced adverse reactions is the frequency with which depressed patients report complaints considered as ADRs prior to receiving any antidepressant medication. This blurs the line between depressive symptoms themselves and unwanted side effects that result from treatment of depression, and represents a complex issue to disentangle. However entering baseline reports of ADRs and depression symptom scores as

covariates into the analysis helps to ensure the associations between genotype and side effects that are seen here are not driven by some factor linked to illness severity.

Arguably, the most important limitation in terms of detecting genetic associations with antidepressant side effects is treatment discontinuation. As has been highlighted, antidepressant discontinuation is strongly associated with the experience of side effects (Bull *et al*, 2002; Mitchell, 2006). Therefore, it is likely that those experiencing more severe side effects were more likely to drop out of research projects early. The general pattern of side effect ratings in GENDEP support this proposal, as ratings were consistently higher in the week prior to drop out in those discontinuing treatment than in those who remained in the study. This non-randomly missing data may bias the findings presented here; however this bias is towards the null, and may result in an underestimate of genotype-ADR associations.

3.5.4 Conclusions

In this chapter, antidepressant ADRs were categorised using knowledge about their shared underlying pharmacological basis. However, when relating the variability seen in the appearance of these pharmacologically-defined side effect groups to genetic polymorphisms within those receptors, predictors of ADRs could not be identified for adrenergic, cholinergic or histaminergic side effects. Significant associations were observed between the occurrence of serotonergic side effects and variability within the *HTR2C* gene in the GENDEP sample, however this finding could not be confirmed using GenPod as a replication dataset.

On expansion of the regions considered surrounding the candidate genes, the *HTR2C* association with serotonergic side effects remained significant. An additional association observed between a single variant in *CHRM2* and cholinergic side effects were seen in the whole sample, but not in drug-specific analyses. This finding could not be tested within the GenPod dataset.

If robust predictors of specific sub-types of antidepressant side effects could be identified (genetic, or otherwise), patients with a high risk of suffering from particular subtypes of side effects could be identified prior to treatment commencement, and treatment options guided accordingly. Whilst the results presented here provide evidence of potential genetic underpinnings of antidepressant ADRs, the absence of replication indicates caution in interpreting these findings is needed.

Chapter 4 CYP450 enzymes and antidepressant treatment response

The work presented here is published:

Hodgson K, Tansey K, Dernovsek MZ, Hauser J, Henigsberg N, Maier W*, et al* (2014). Genetic differences in cytochrome P450 enzymes and antidepressant treatment response. *Journal of Psychopharmacology* **28**(2): 133-141.

4.1 Introduction

4.1.1 Role of CYP450 enzymes in the metabolism of antidepressants

CYP450 enzymes are involved in the metabolism of 75% of marketed drugs (Guengerich, 2008). As outlined in the Introduction, a number of common polymorphisms in the genes encoding the CYP450 enzymes are associated with differences in enzyme activity. Given that variability in the rate of drug metabolism could alter the concentrations of active compound, it has been suggested that genotyping the CYP450 enzymes may allow many drug therapies to be "personalised" by matching drug doses to an individual's genetically predicted drug metabolism rate (Ingelman-Sundberg, 2004).

Antidepressants are amongst those drugs that are metabolised by CYP450 enzymes. With reference to those antidepressants studied in GENDEP, the tricyclic antidepressant nortriptyline is metabolised to both *cis*- and *trans*- forms of 10-hydroxynortriptyline via the CYP2D6 enzyme (Olesen *et al*, 1997a). In contrast, three CYP450 enzymes are involved in the metabolism of escitalopram to its primary metabolite desmethylescitalopram; CYP3A4, CYP2C19 and, to a lesser extent, CYP2D6 (Olesen and Linnet, 1999).

4.1.2 Impact of genetic variability in CYP450 enzymes

Genetic variation in the *CYP3A4* gene has been reported to be rare and have little impact on enzymatic activity (Lamba *et al*, 2002). However, genetic differences in *CYP2C19* and *CYP2D6* do have functional consequences on enzymatic activity (De Morais et al, 1994; Eichelbaum et al, 1979; Ferguson et al, 1998; Mahgoub et al, 1977; Tucker et al, 1977).

Previous reports have found that the genetic variation in *CYP2C19* is related to serum concentrations of escitalopram in patients being treated with the antidepressant; this has been found both within a subset of the GENDEP sample (Huezo-Diaz *et al*, 2012) and in other datasets (Rudberg *et al*, 2006; Rudberg *et al*, 2008). Similarly nortriptyline concentrations have been associated with polymorphisms in *CYP2D6*, albeit in smaller sample sizes (Dahl *et al*, 1996; Murphy *et al*, 2001).

Whilst previous literature indicates there is a relationship between CYP450 genotypes and antidepressant serum concentrations, this has not been translated robustly into clinically important differences in treatment response. Kirchheiner and colleagues reviewed the relationship between CYP450 genotypes in *CYP2D6* and *CYP2C19* and serum concentration of drug, attempting to establish genotype-specific dose adjustments for a range of antidepressants, but they did not examine the effect on treatment outcome (Kirchheiner *et al*, 2001; Kirchheiner *et al*, 2004). More recent studies into the role of CYP450 genetic variation on antidepressant treatment response report discrepant findings (Mrazek *et al*, 2011; Peters *et al*, 2008; Tsai *et al*, 2010). The results of a pilot study (Hall-Flavin *et al*, 2012) suggested outcomes could be improved amongst depressed patients when treatment is directed by a pharmacogenomics algorithm which included information on cytochrome P450 genotypes. However the sample size in this pilot study was small, and the information contained within the algorithm is not limited to drug metabolising genes, but also encompasses genes involved in the serotonergic system. To date, there remains a lack of direct evidence for a relationship between CYP450 genotype and differences in antidepressant treatment outcomes.

4.1.3 Serum concentrations of antidepressant

Whilst CYP450 genotype is observed to play a role in determining serum concentration of antidepressant, other factors such as diet, comedication and comorbidities are also known to contribute to the large inter-individual variability observed in serum concentrations (Reis *et al*, 2009). Therefore, measuring serum levels of antidepressant allows us to consider both genetic and environmental influences together in order to understand the role that drug metabolism differences may have on treatment response.

Previous work reporting a curvilinear relationship between peripheral concentrations of circulating nortriptyline and treatment response may indicate a complex relationship between serum drug levels and treatment response (Asberg *et al*, 1971; Perry *et al*, 1994). But when investigating citalopram (the racemic form of escitalopram), evidence indicates there is not a link between serum concentrations of antidepressant and likelihood of treatment response (Dufour *et al*, 1987; Nikisch *et al*, 2004; Rasmussen *et al*, 2000). Nevertheless, a recent study observed that patients with citalopram levels above 50µg/L had improved scores of depression severity than those with levels below this concentration (Ostad Haji *et al*, 2011).

The Arbeitsgemeinschaft für Neuropsychopharmakologie und Pharmakopsychiatrie (AGNP) Consensus Guidelines for 2011 (Hiemke *et al*, 2011) recommend therapeutic dose monitoring for both nortriptyline and citalopram. However, particularly in the case of citalopram, this is based on evidence predominantly related to the observed variability in the pharmacokinetics of the drug, rather than direct tests of the relationship with serum levels of the drug and treatment outcomes.

4.1.4 Aims

This chapter aims to directly test the relationship between genetic variation in CYP450 enzymes, serum concentration of antidepressants and response to treatment. Using the data available in GENDEP, the three objectives of this chapter were to: (1) to replicate previous work showing that CYP450 genotype influences serum concentrations of antidepressant; (2) to examine whether CYP450 genotype predicted antidepressant treatment response; and (3) to explore the relationship between serum concentrations of antidepressant and treatment response.

4.2 Methods

4.2.1 Participants

Participants included in this chapter are drawn from the GENDEP sample, as described in detail in Chapter 2.2. Figure 4-1 shows the patients included in analyses presented in this chapter.

4.2.2 CYP450 Genotypic information

As described in Chapter 2.6.1, genotyping of the CYP450 enzymes *CYP2C19* and *CYP2D6* was achieved using the Roche AmpliChip CYP450 (Roche Molecular Diagnostics, Alameda, CA, USA). Functional annotations were made for each of the alleles contained on the microarray using information as described by the Human Cytochrome P450 (CYP) Allele Nomenclature Database (http://www.cypalleles.ki.se/). Allele frequencies across the entire GENDEP sample are reported in Table 4-1 and Table 4-2.

After annotations were assigned to each allele, genotypic classification was performed. We used four categories to classify *CYP2D6* genotypes; poor (PM), intermediate (IM), extensive (EM) or ultrarapid (UM) metabolisers (Rebsamen *et al*, 2009). For *CYP2C19*, we used six categories; poor (PM), intermediate (IM), intermediate plus (IM+), extensive (EM), extensive plus (EM+) or ultrarapid (UM) metabolisers (Mrazek *et al*, 2011). The allele frequencies in each of the drug-specific subsets of GENDEP are shown in Table 4-3 and Table 4-4.

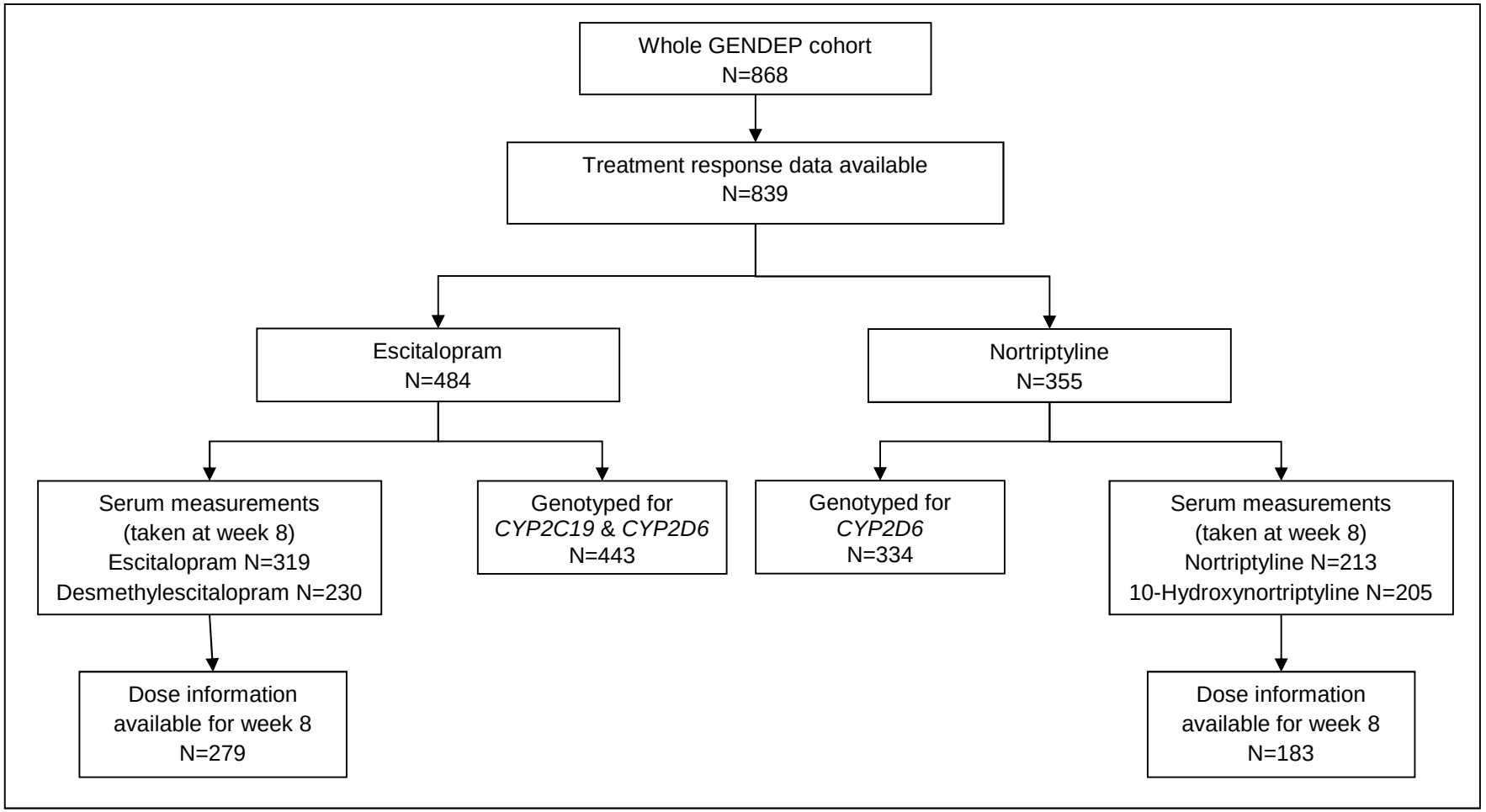

Figure 4-1: Sample included in Chapter 4

Table 4-1: Functional annotations for alleles identified in *CYP2D6*

Functional effects	Allele name (CYP nomenclature)	Frequency in GENDEP sample
Absent	*3	30
	*4	310
	*5	63
	*6	22
	*7	4
	*8	-
	*11	-
	*14A	-
	*15	2
	*19	-
	*20	-
	*40	-
	*4 x N	3
Decreased	*9	30
	*10	25
	*17	4
	*29	-
	*36	-
	*41	137
	*10 x N	-
	*17 x N	-
	*41 x N	3
Normal	*1	572
	*2	255
	*35	112
Increased	*1 x N	13
	*2 x N	17
	*35 x N	2
Unknown	*14B	1
	*25	1
	*26	-
	*30	-
	*31	-

Table 4-2: Functional annotations for alleles identified in *CYP2C19*

Functional consequence	Allele name (CYP nomenclature)	Frequency in GENDEP sample
Decreased	*2	249
Normal	*1	999
Increased	*17	394

Table 4-3: Frequencies of genotypic categories for CYP2D6 amongst patients taking nortriptyline. (Total N=334)

Genotypic group	Classification criteria	Frequency in GENDEP sample	
		N	%
PM	2 non-functional alleles	23	6.89
IM	1 non-functional, 1 decreased / 2 decreased alleles	23	6.89
EM	At least 1 wild type allele	281	84.13
UM	At least 3 copies of a wild type allele	7	2.10

Table 4-4: Frequency of genotypic categories for *CYP2C19* and *CYP2D6* amongst patients taking escitalopram. (Total N=443)

Genotypic group	Classification criteria	Frequency in GENDEP sample	
		N	%
	CYP2C19		
PM	*2/*2	6	1.35
IM	*1/*2	85	19.19
IM+	*17/*2	34	7.67
EM	*1/*1	176	39.73
EM+	*17/*1	115	25.96
UM	*17/*17	27	6.09
	CYP2D6		
PM	2 non-functional alleles	34	7.67
IM	1 non-functional, 1 decreased / 2 decreased alleles	37	8.35
EM	At least 1 wild type allele	354	79.91
UM	At least 3 copies of a wild type allele	18	4.06

4.2.3 Serum concentrations of antidepressant

As described in Chapter 2.5, concentrations of antidepressant (escitalopram/nortriptyline) and primary metabolite (desmethylescitalopram/total 10-hydroxynortriptyline) were measured in blood samples taken at week eight of antidepressant treatment. For ease of interpretation, standardised serum measurements were calculated, with a mean of 0 and a standard deviation of 1. The ratio of metabolite to drug concentration was also calculated.

Daily prescribed dose of antidepressant was recorded. Of those with serum measurements, 86.84% had information available on the dose of antidepressant taken that week. For additional analyses, dose-adjusted serum concentrations were calculated.

4.2.4 Comedications

As described in Chapter 2.3.1, patients reported all medications taken throughout the study, in addition to the prescribed antidepressant. These comedications were categorised using the FDA classification of *in vivo* inhibitors or inducers (FDA, 2011)http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DevelopmentResources/DrugInteractionsLabeling/ucm093664.htm%23classInhibit. In week eight (the week that blood samples were taken for serum concentration measurements), 4.51% of patients with serum concentration data reported taking a CYP2C19 and/or CYP2D6 inhibiting drug, but no patients reporting taking CYP2C19 and/or CYP2D6 inducing drugs. The specific drugs reported are shown in Table 4-5 and Table 4-6.

Table 4-5: CYP450-inhibiting comedication reported by patients prescribed escitalopram (N=319).

Drug taken	CYP450 Enzymes Inhibited	Number of patients
Combined Oral Contraceptive Pill	CYP2D6 and CYP2C19	11
Diltiazem	CYP2D6	1
Propafenone	CYP2D6	1
Verapamil	CYP2D6	1
Omeprazole	CYP2C19	1

Table 4-6: CYP2D6-inhibiting medication reported by patients prescribed nortriptyline (N=213)

Drug taken	CYP450 Enzymes Inhibited	Number of patients
Combined Oral Contraceptive Pill	CYP2D6 and CYP2C19	8
Amiodarone	CYP2D6	1
Ranitidine	CYP2D6	1

4.2.5 Treatment response

Treatment response was measured using the MADRS, as detailed in Chapter 2.4.1. All weekly ratings of depression severity were used within a repeated measures design.

4.2.6 Statistical analysis

All analyses were performed using STATA 11 (StataCorp., 2009). Given the different metabolic pathways of nortriptyline and escitalopram, all analyses were performed in a drug-specific manner.

For escitalopram**,** *CYP2C19* was examined as the relevant CYP450 genotype and we included *CYP2D6* genotype as a fixed effect covariate given the smaller reported role of the CYP2D6 enzyme (Olesen *et al*, 1999). For nortriptyline, *CYP2D6* was used as the relevant CYP450 genotype in the nortriptyline-specific analysis.

To account for the clustering of individuals around recruitment centres, mixed models were used (fitted with maximum likelihood); centre was entered as a random effect. Fixed effects included age and sex.

4.2.6.1 CYP450 genotype predicting serum concentrations of antidepressant

The first stage of the analysis aimed to examine CYP450 genotype as a potential predictor of three standardised serum measures; concentrations of the antidepressant, its primary metabolite, and the metabolite to drug ratio. Cytochrome P450-inhibiting comedications were included as a fixed effect in the model.

4.2.6.2 CYP450 genotype predicting treatment response

The second stage of the analysis aimed to examine CYP450 genotype as a potential predictor of treatment response (using weekly measures as described above). Fixed effects of baseline depression severity, cytochrome P450-inhibiting comedications and both linear and quadratic effects of time were included. Given the longitudinal analysis of response, individual was included as a random effect (in addition to recruitment centre).

4.2.6.3 Serum concentration of antidepressant predicting treatment response

The final stage of the analysis considered whether serum concentrations of antidepressant were predictive of treatment response. Three potential predictors were explored: concentrations of drug, primary metabolite and ratio of metabolite to drug. For any significant associations, the analysis was repeated, with dose included as a fixed effect. This is particularly important in light of the flexible dosing protocol used in GENDEP, where daily drug dose could be altered over

the course of the study if necessary (for example, in response to limited improvement in depressive symptoms).

Given previous reports of a curvilinear association between serum concentration and response to treatment with nortriptyline (Asberg et al., 1971), an additional analysis of the relationship between serum concentration of antidepressant and treatment response was also performed, where serum measurements were centred and then squared.

4.3 Results

4.3.1 Descriptive statistics

4.3.1.1 CYP450 genotype

The observed CYP450 genotypic frequencies match those expected in Caucasian populations, with the most common genotypic group being EM (extensive metaboliser) for both CYP450 enzymes (de Leon *et al*, 2006; Mrazek *et al*, 2011; Rebsamen *et al*, 2009). Genotype was unrelated to recruitment centre, baseline depression severity, age, dose or sex, in either those taking escitalopram or nortriptyline.

4.3.1.2 Serum concentration of antidepressant

4.3.1.2.1 Serum concentrations and dose of antidepressant prescribed

The mean serum concentrations of antidepressant are shown in Table 4-7, along with the prescribed daily dose of antidepressant for week eight of the study. Serum levels are comparable to those observed in other studies (Foglia *et al*, 1997; Vandel *et al*, 1978).

Table 4-7: Daily prescribed dose and serum concentration of drug

Drug taken		N	Mean	SD
Escitalopram	Daily dose (mg)	279	15.99	6.40
	Serum escitalopram (µg/L)	319	29.04	18.40
	Serum desmethylescitalopram (µg/L)	230	10.99	5.07
Nortriptyline	Daily dose (mg)	183	103.66	32.25
	Serum nortriptyline (µg/L)	213	91.07	57.99
	Serum total 10-hydroxy-nortriptyline (µg/L)	205	65.74	51.54

Concentrations of drugs and primary metabolites were unrelated to age in this sample. Females had significant higher concentration of both escitalopram (female; M=31.36, SD=19.30, male; M=24.50, SD=19.29; t(317)=-3.20, p=0.0015) and desmethylescitalopram (female; M=11.46, SD=5.28, male; M=9.88, SD=4.34, t(228)=-2.16, p=0.032). No significant differences were observed between males and females for nortriptyline and 10-hydroxynortriptyline concentrations. As expected, dose was significantly related to serum levels of antidepressant measured at week eight (see Figure 4-2).

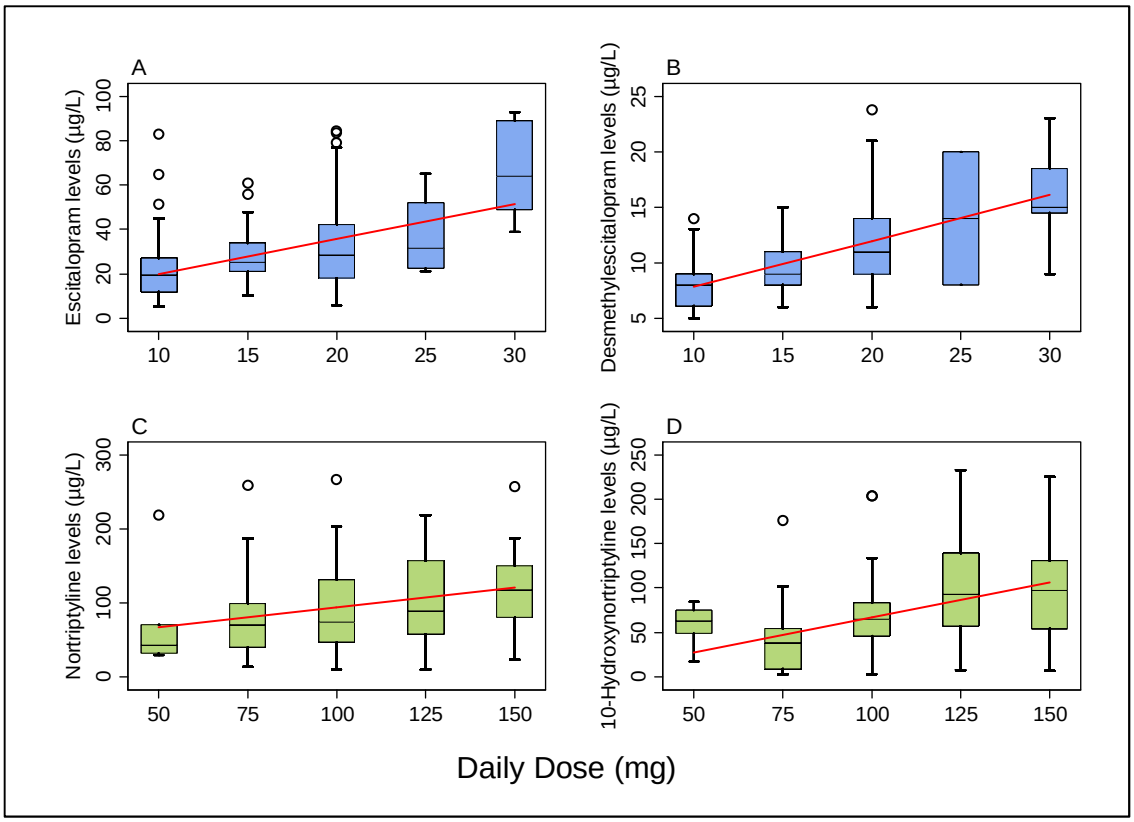

Figure 4-2: Relationship between dose and serum level of antidepressant.
All associations are highly significant (p<0.0001). Proportion of variance explained by dose in each of the analyses; A) 0.29 B) 0.36 C) 0.08 D) 0.21.

There were also differences between centre of recruitment in terms of serum concentrations of antidepressant, daily dose prescribed and also in dose-adjusted serum concentrations of antidepressant and metabolite (Table 4-8). This emphasises the importance of including recruitment centre as a random effect in all models.

Table 4-8: Variation between recruitment centres in dose and dose-adjusted serum concentration in GENDEP

Drug taken	Analysis	F-statistic	p-value
Escitalopram	Daily dosage	(8,270)=14.70	8.00×10^{-18}
	Dose-adjusted escitalopram level	(8,248)=3.45	0.0009
	Dose-adjusted desmethylescitalopram level	(8,188)=2.34	0.0201
Nortriptyline	Daily dosage	(8,174)=10.67	3.76×10^{-12}
	Dose-adjusted nortriptyline level	(8,160)=3.04	0.0033
	Dose-adjusted 10-hydroxynortriptyline level	(7,158)=6.40	1.23×10^{-6}

4.3.1.2.2 CYP450-inhibiting comedication and serum concentrations

All the CYP450-inhibiting comedications reported (see Table 4-5 and Table 4-6) were classified by the FDA as “weak” inhibitors of CYP2D6 or CYP2C19. Amongst patients on escitalopram, those taking CYP2C19-inhibiting comedications had significantly higher concentrations of

escitalopram (M=40.41, SD=18.04) than those not taking these medications (M=28.51, SD=18.27; t(317)=-2.38, p=0.0178). These comedications did not affect desmethylescitalopram concentrations. Patients taking nortriptyline along with CYP2D6-inhibiting comedications had higher concentrations of nortriptyline (M=145, SD=45.14) than those without comedications (M=88.69, SD=57.42; t(211)=-2.90, p=0.0041). They also had significantly higher levels of total 10-hydroxnortriptyline (M=102.33, SD=58.82,) than those without comedications (M=64.06, SD=50.72; t(203)=-2.20, p=0.0290). All analyses included CYP450-inhibiting comedications as a covariate. Additionally, when those individuals taking CYP450-inhibiting comedications are excluded from the analysis, the pattern of findings remained unchanged.

4.3.1.2.3 Comparison of participants with and without serum measurement data available

As can be seen from Figure 4-1, serum measurements were not available for all participants. Therefore, differences between those participants where serum measurements were available and those where this data was missing were explored.

There were no observed differences between those with and without serum measurements in age, sex, baseline depression severity or CYP450 genotype.

However, significant differences were observed in terms of treatment response outcomes. Of those without serum measurements, 52% had dropped out of the study prior to week eight of the trial. Nevertheless, a comparison of treatment response as measured by adjusted percentage change in MADRS scores (Uher *et al*, 2010) in those cases who had remained in the study until at least week eight found significant differences (t(653)=3.13, p=0.0018); those with serum measurements available (M=4.40, SD=27.39) were more likely to have responded to treatment than those where serum measurements were unavailable (M=-4.13, SD=33.83).

4.3.2 CYP450 predicting serum concentrations of antidepressant

4.3.2.1 Escitalopram

Genetic variation in *CYP2C19* significantly predicted serum concentration of escitalopram (n=264, ß=-0.232 SE=0.044, $p=1.06x10^{-7}$); genotypes that encode more active forms of the CYP2C19 enzyme were associated with lower levels of escitalopram. *CYP2C19* genotype was not associated with concentration of desmethylescitalopram (n=196, ß=0.085 SE=0.048, p=0.072), but was significantly associated with the ratio of desmethylescitalopram to escitalopram (n=195, ß=0.281, SE=0.054, $p=2.31x10^{-7}$); higher levels of enzyme activity were associated with higher metabolite to drug ratios (Figure 4-3).

4.3.2.2 Nortriptyline

Genetic variation in *CYP2D6* significantly predicted dose-adjusted levels of both nortriptyline (n=173, ß=-0.562 SE=0.107, $p=1.64x10^{-7}$) and total 10-hydroxy-nortriptyline (n=168, ß=0.327, SE=0.106, p=0.0021). Genotypes encoding more active forms of the *CYP2D6* enzyme were linked to lower serum concentrations of nortriptyline and higher concentrations of total 10-hydroxy-nortriptyline. Genetic variation linked to higher levels of *CYP2D6* activity were associated with higher 10-hydroxy-nortriptyline to nortriptyline ratios (n=167, ß=0.496, SE=0.114, $p=1.37x10^{-5}$, Figure 4-4)

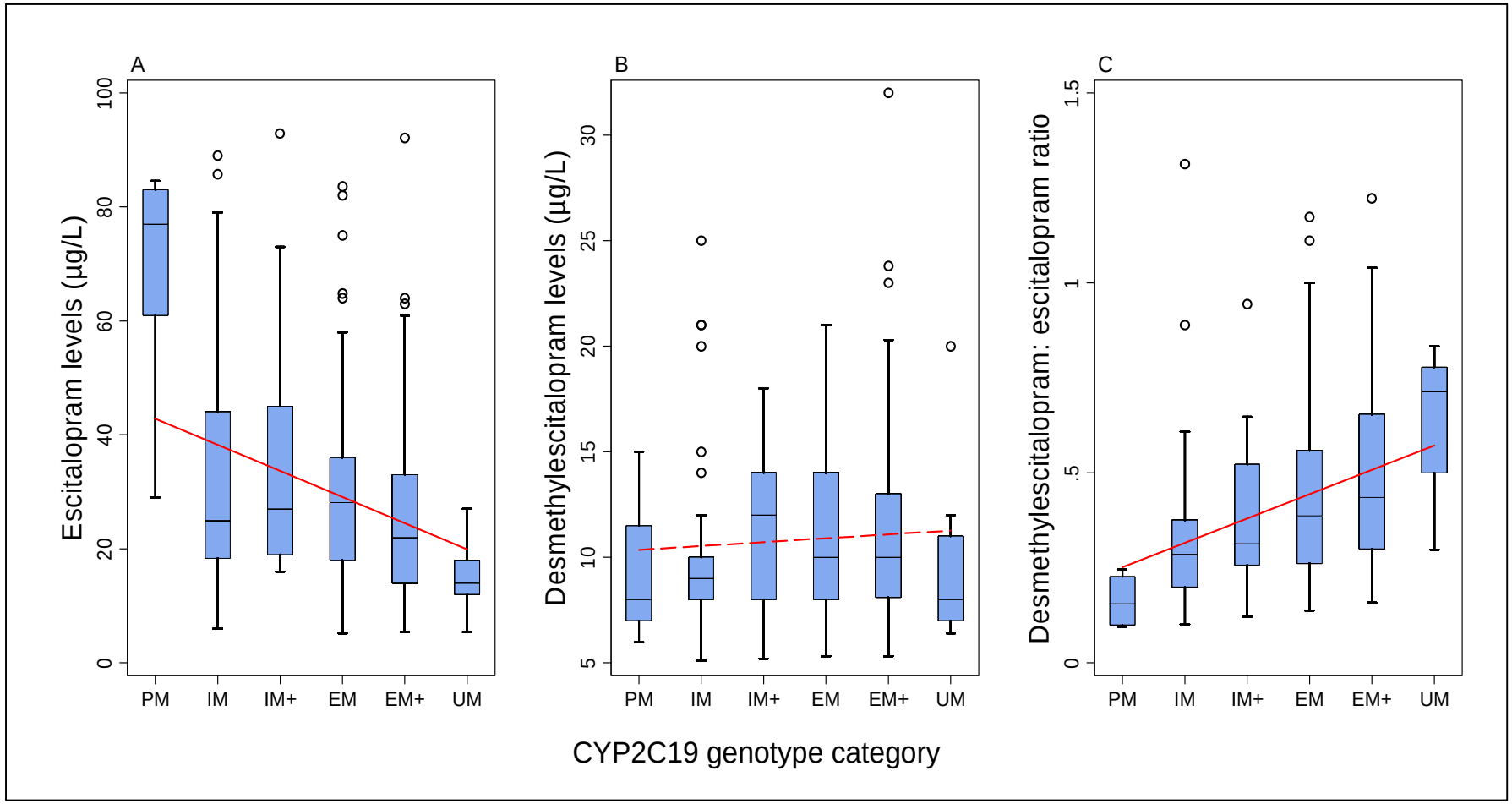

Figure 4-3: Relationship between *CYP2C19* genotype and serum concentration of escitalopram.

(Dashed red line in panel B indicates non-signficant relationship)

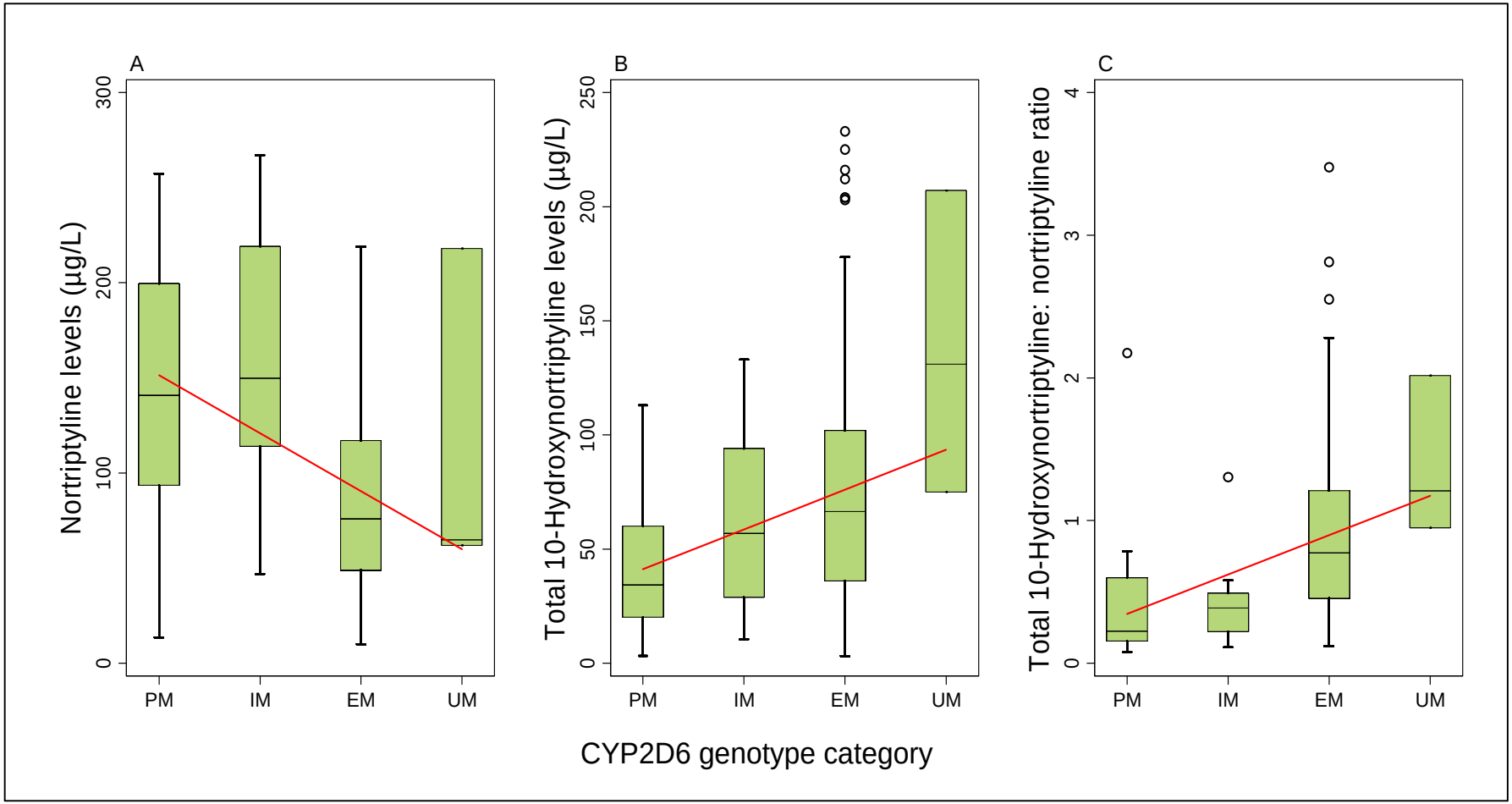

Figure 4-4: Relationship between *CYP2D6* genotype and serum concentration of nortriptyline.

4.3.3 CYP450 genotype and treatment response

There was no significant relationship between genetic variation in CYP450 enzymes and response to treatment with antidepressants in either of the two medications examined in this study (escitalopram, n=443, p=0.478, ß=0.165, SE= 0.233; nortriptyline, n=334, p=0.807, ß=0.127, SE=0.524).

4.3.4 Serum concentration of antidepressant and treatment response

For patients taking escitalopram, there was a significant association between serum concentrations of escitalopram and treatment response (n=319, p=0.004, ß=0.932, SE=0.321). However, this association was not in the expected direction, with higher serum concentrations of escitalopram predicting poorer treatment response. When drug dose was added as a covariate, this association no longer reached significance (p=0.207). Neither desmethylescitalopram concentrations nor the ratio of desmethylescitalopram to escitalopram significantly predicted response.

In those patients taking nortriptyline, higher concentrations of total 10-hydroxy-nortriptyline were significantly linked to poorer treatment response (n=205, p=0.006, ß=1.144, SE=0.416). Again, this association no longer reached significance when dose was entered as a covariate (p=0.183). There was no significant association between treatment response and with concentration of nortriptyline or metabolite to drug ratio.

There was no evidence of a curvilinear relationship between treatment response and serum concentration of drug, primary metabolite or metabolite to drug ratio for either drug.

4.4 Discussion

4.4.1 Summary of results

4.4.1.1 CYP450 genotype

This analysis has shown that CYP450 genotype predicts serum concentration of antidepressant for patients taking either escitalopram or nortriptyline. However, there is no relationship between CYP450 genotype and treatment response for the patients in GENDEP.

4.4.1.2 Serum concentration of antidepressant

Contrary to expectations, higher concentrations of drug predicted poorer treatment response for patients taking escitalopram, and higher concentrations of metabolite predicted poorer treatment response for patients taking nortriptyline. However, controlling for the effects of drug dosage removes these effects, indicating that it is likely that the relationship between serum concentration and poor response is due to the dosing protocol used in GENDEP. Serum samples were taken after eight weeks of treatment, during which time clinicians were able to follow a protocol-driven flexible dosing schedule, informed by depressive symptoms and side effects. Thus, higher drug doses were prescribed to those patients who were failing to adequately respond to treatment. No evidence of a curvilinear association between serum concentration and treatment response was observed for either antidepressant.

4.4.2 Comparison to previous literature

4.4.2.1 CYP450 genotype

The evidence presented here does not support early hopes that pharmacogenetic examination of pharmakinetic processes would prove clinically useful (Ingelman-Sundberg, 2004). Whilst the association between CYP450 genotype and serum concentration is as previously reported (Dahl *et al*, 1996; Huezo-Diaz *et al*, 2012; Kirchheiner *et al*, 2001; Kirchheiner *et al*, 2004; Murphy *et al*, 2001; Rudberg *et al*, 2008), this association cannot be translated into differences in treatment response.

This is in line with a report from the US-based antidepressant study STAR*D, where response to treatment with citalopram was assessed in relationship to *CYP2C19* genotype in a sample of nearly 2,000 subjects (Peters *et al*, 2008). Nevertheless, a second paper on the same sample drew contradictory conclusions, reporting some evidence of a link between CYP450 genotype and response (Mrazek *et al*, 2011). In this analysis, we have not only replicated the finding of Peters et al, but also extended this to include the tricyclic antidepressant nortriptyline.

4.4.2.2 Serum concentration

Furthermore, GENDEP is the largest study to date to incorporate measurements of antidepressant serum concentrations and together with comprehensive CYP450 genotypic information. Our failure to find links between serum concentration and treatment response seems to be consistent with the majority of studies looking at citalopram (Dufour *et al*, 1987; Nikisch *et al*, 2004; Rasmussen *et al*, 2000), but is in contrast to classical work looking at serum levels for nortriptyline (Asberg *et al*, 1971; Perry *et al*, 1994), which indicated there may be a curvilinear association between serum concentration and treatment response.

4.4.3 Methodological considerations

4.4.3.1 Power calculation

In light of the negative associations reported here, a post-hoc power calculation was undertaken to establish the statistical power that this sample has to detect clinically meaningful effects of drug metabolism enzymes. Uher et al. (2012) calculated that for studies addressing predictors of antidepressant treatment outcomes, continuous biomarkers (such as serum levels in this study) should explain at least 6.3% of the variance in treatment response in order to be clinically significant. Using G*Power (http://www.psycho.uni-duesseldorf.de/abteilungen/aap/gpower3/download-and-register), it was calculated that a sample size of n=120 would be needed to detect an effect size of this magnitude with p=0.05, and power of 80%. Each analysis presented here exceeds that sample size, so is adequately powered to detect clinically significant associations between serum levels of antidepressant and treatment response. Thus the failure to observe an association with treatment response cannot be attributed to a lack of statistical power.

4.4.3.2 Multiple-hypothesis testing

No formal multiple-hypothesis testing corrections were applied to control for false positive associations, in light of the interdependence of the measures explored here. Nonetheless, the key significant findings of CYP450 genotype being associated with antidepressant concentration are highly significant (escitalopram; $p=1.06x10^{-7}$, nortriptyline; $p=1.64x10^{-7}$), and so are unlikely to be false positives. Neither CYP450 genotype nor dose-adjusted serum levels predicted treatment response, even at the uncorrected cut-off for statistical significance of $p<0.05$.

4.4.3.3 CYP450 genotypic categorisation

There remains some debate over the most appropriate model to use when categorising CYP450 genotypes. CYP2C19 genotypic variation has been categorised into six subgroups (Mrazek *et al*, 2011), four subgroups (Brandl *et al*, 2014), or two categories (Peters *et al*, 2008); in this study the six category model was used, as by Mrazek et al., because this model utilises the full information obtained by genotyping the more recently reported *17 allele (Sim *et al*, 2006). Similarly, CYP2D6 can be considered within a model of 4 genotypic categories (Brandl *et al*, 2014; Rebsamen *et al*, 2009) or two categories (Peters *et al*, 2008), here the 4 category model was used as the more comprehensive option. However, the pattern of results reported here, where CYP450 genotype predicts dose-adjusted serum levels of antidepressants but not treatment response, remain robust to the use of alternative models of CYP450 genotype grouping.

4.4.4 Study limitations

Despite the robustness of the findings to these methodological considerations, there remain limitations to the study.

4.4.4.1 Treatment response biases

Whilst there are no significant differences in baseline characteristics and CYP450 genotype between patients with and without serum data, in those patients where the relevant data is available, treatment response is significantly better amongst those with serum measurements.

This may be a source of bias in the study. The reasons for this are not known, but potentially, those patients with greater improvement in depression symptoms may be more likely to be fully engaged in the study, and so consent to the additional blood sample taken at week eight of the study.

4.4.4.2 Recruitment centre variation

Significant differences in serum concentration are observed between recruitment centres, even after dose-adjustments are made. Given that CYP450 genotype shows no association with recruitment centre, these variations may be due to environmental differences between populations recruited to the different centres across Europe; for example diet is known to affect drug metabolism rates (Conney *et al*, 1980). Whilst recruitment centre is added as a random effect in all analyses, these differences add variability to the data.

4.4.4.3 Flexible dosing protocol

Finally, the flexible dosing protocol used in GENDEP means that dosage was titrated according to clinical need (in response to non-response or the occurrence of side effects). Therefore, our conclusions indicate the extent to which data on CYP450 genotypes and serum concentrations of antidepressants are predictive of treatment response when used in addition to clinical observation, but cannot be extended beyond this.

4.4.5 Conclusions

CYP450 enzymes play a key role in the metabolism of antidepressants, and the efficacy of these enzymes varies across the population as a result of genotypic differences. This chapter has shown that these genetic factors have a significant impact on serum concentrations of antidepressants in this sample. However, CYP450 genotype could not be used as a predictor of treatment response and there was no clear relationship between serum concentrations of antidepressant and differences in treatment response in this study. Therefore, within the context of clinical observation and a flexible dosing protocol, drug metabolism variability appears to be

uninformative regarding treatment response. In the next chapter, the effects of drug metabolism variability on the outcomes of side effects and study drop out are explored.

Chapter 5 CYP450 enzymes and antidepressant side effects

5.1 Introduction

5.1.1 The importance of side effects as an antidepressant treatment outcome

As noted in Chapter 3, antidepressant medications are associated with a number of unwanted side effects (or adverse drug reactions) including dry mouth, sexual dysfunction and gastrointestinal symptoms. Whilst selective serotonin reuptake inhibitors (SSRIs) were developed and marketed as drugs with improved tolerability compared to the older tricyclic antidepressants (Whyte *et al*), and indeed are safer in overdose, side effects with these newer drugs are still high. Estimates vary, but the frequency of adverse drug reaction (ADRs) with SSRI treatment has been observed to be as high as 75% (Meijer *et al*, 2002).

Side effects play an important role in clinical decision-making when prescribing antidepressants; it is known that the experience of adverse drug reactions is strongly associated with discontinuation and poor treatment adherence (Bull *et al*, 2002; Mitchell, 2006). As would be expected, inadequate antidepressant treatment results in increased levels of relapse (Kennedy *et al*, 2002) as well as increased health costs (Masand, 2003). Therefore, the ability to predict which individuals are at higher risk of experiencing side effects or discontinuing treatment would be useful in a clinical setting.

5.1.2 Role of drug metabolism rates in risk for side effects

The value of exploring adverse drug reactions in relation to drug metabolism variability is underscored by examples within other medical areas. For example, for the anticoagulant warfarin, the observed association between drug metabolism rates and adverse drug reactions has led to therapeutic dose monitoring being used as standard within NHS treatment guidelines, and extensive research into the viability of translating a pharmacogenetic test for clinical use (Gage *et al*, 2008; Kangelaris *et al*, 2009).

Considering adverse drug reactions to antidepressant treatment, the Clinical Pharmacogenetics Implementation Consortium issued guidelines in 2013 regarding dosing of tricyclic antidepressants (Hicks *et al*, 2013). With reference to nortriptyline, the consortium concluding

that there was strong evidence to avoid nortriptyline or amitriptyline (which is metabolised in the body to nortriptyline) amongst CYP2D6 poor metabolisers due to potential for side effects. However, much of this evidence is based on the observed relationship between CYP450 genotype and serum concentrations of antidepressant; there is a paucity of evidence evaluating whether CYP450 genotype is associated with treatment outcomes in patient samples. Nevertheless, the guidelines do reference two studies which observe associations between drug metabolising variables and measures of adverse reactions to tricyclic antidepressants (Bijl *et al*, 2008; Steimer *et al*, 2005).

With reference to SSRIs, a comprehensive review in 2007 found six studies exploring the relationship between CYP450 genotype and side effects, four of which reported no association. However, the authors of the review highlighted the limited power of these studies and concluded that no firm conclusions could be drawn regarding the association between CYP450 genotype and SSRI-associated ADRs.

Since this review was published, the role of genetic variation in the CYP450 enzymes for patients taking SSRIs has been further explored in the STAR*D study, with two reports exploring treatment tolerance alongside response measures (Mrazek *et al*, 2011; Peters *et al*, 2008). Treatment tolerance in STAR*D was assessed using study exit data and patient decisions on whether to continue with citalopram treatment. Mirroring the findings in treatment response (detailed in Chapter 4), whilst one report found no evidence of association with tolerance amongst patients taking citalopram (Peters *et al*, 2008), a more recent paper based on the same data concluded that there was evidence of a link between CYP450 genotype and tolerance (Mrazek *et al*, 2011).

5.1.3 Aims

In light of the importance of adverse drug reactions in determining outcomes to antidepressant treatment, and the lack of clear evidence demonstrating links between drug metabolism rates

and ADRs, this chapter aims to extend the work of Chapter 4, examining the potential links between rates of drug metabolism and antidepressant side effects.

The in-depth data available in the GENDEP sample allows a detailed analysis of side effects reported with antidepressant treatment. Firstly side effect burden was considered as a quantitative trait, by totalling the number of side effects reported, per individual, per week. Secondly, the relationship between drug metabolism and each specific ADR was explored. This approach allows consideration of the differences between side effects in terms of their pharmacological basis.

As with the analysis in Chapter 4, drug metabolism rates were considering in two ways; firstly using genotypic information from the CYP450 enzymes of relevance, and secondly using serum concentrations of drug, which allow us to capture the combined effects of both genetic and environmental influences on drug metabolism rates.

Finally, we explored the association between CYP450 genotype and study drop out, to enable comparison with previous literature and address general tolerability of treatment.

5.2 Methods

5.2.1 Participants

Participants included in this chapter are drawn from the GENDEP sample, as described in detail in Chapter 2.2. Figure 5-1 shows the patients included in analyses presented in this chapter.

5.2.2 CYP450 genotypic information

The genotyping of CYP2C19 and CYP2D6, using the Roche AmpliChip is described in detail in Chapter 2.6.1, with the functional annotations and genotypic frequencies detailed in Chapter 4. In brief, four genotypic categories were used to classify *CYP2D6* genotypes (PM, IM, EM and UM; (Rebsamen *et al*, 2009), whilst *CYP2C19* was classified into 6 genotypic categories (PM, IM, IM+, EM, EM+, UM (Mrazek *et al*, 2011).

5.2.3 Serum concentrations of antidepressant

Serum concentrations of drug (escitalopram or nortriptyline) and primary metabolite (desmethylescitalopram or total 10-hydroxynortriptline) were measured in blood samples taken at week eight of treatment, using achiral turbulent flow liquid chromatography as outlined in Chapter 2.5. These values were standardised, with a mean of 0 and a standard deviation of 1. The ratio of metabolite to drug concentration, and total concentration of drug plus metabolite were also calculated.

5.2.4 Comedications

Reported comedications that may act as CYP450 enzyme inducers or inhibitors were identified using FDA guidelines (FDA, 2011). In the week that we took the blood sample for the measurement of serum levels, 5.81% of patients with data on both serum concentration of antidepressant and side effects reported taking a CYP2C19 and/or CYP2D6 inhibiting drug, but no patients reporting taking CYP2C19 and/or CYP2D6 inducing drugs.

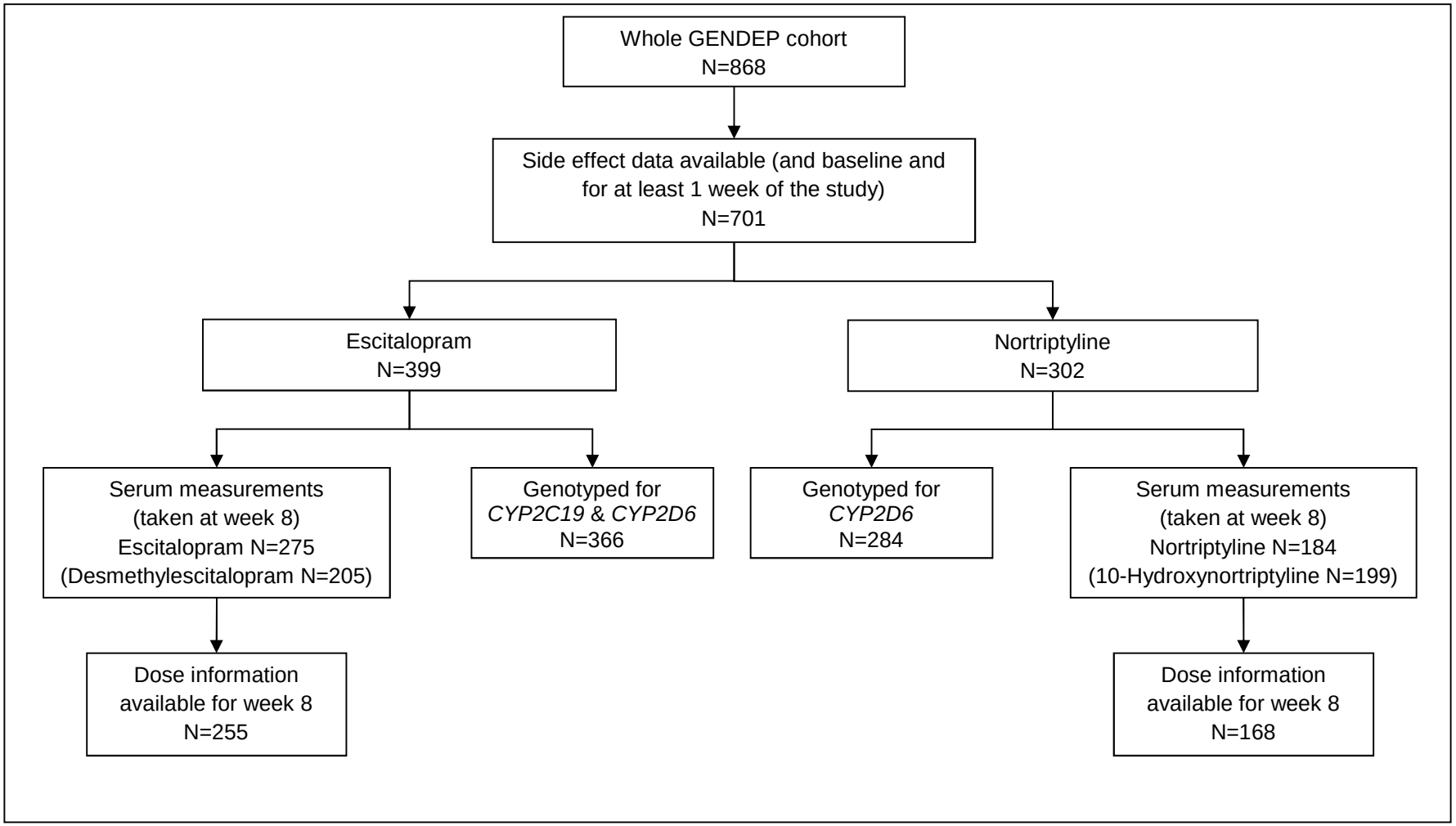

Figure 5-1: Sample included in Chapter 5

5.2.5 Side effects

Antidepressant side effects were measured weekly using the ASEC (see Chapter 2.4.2). Side effect scores were compiled from the ASEC in two ways. Firstly, the total number of ADRs experienced each week was calculated, to give a measure of the overall side effect burden for the patient. Secondly, the weekly presence or absence of each individual side effect was examined separately, to investigate ADR-specific effects.

5.2.5.1 Study drop out

Study drop out was also considered, using the recorded study end week for each participant (detailed in Chapter 2.4.3).

5.2.6 Statistical analysis

The analysis of side effects followed the same framework as described in Chapter 4, when exploring the outcome of treatment response. All analyses were performed using STATA 11 (StataCorp., 2009). Given the different metabolic pathways of nortriptyline and escitalopram, all analyses were performed in a drug-specific manner.

For escitalopram, *CYP2C19* was examined as the relevant CYP450 genotype and we included *CYP2D6* genotype as a fixed effect covariate given the smaller reported role of the CYP2D6 enzyme (Olesen et al, 1999). For nortriptyline, *CYP2D6* was used as the relevant CYP450 genotype in the nortriptyline-specific analysis.

5.2.6.1 CYP450 genotype predicting side effects

To assess the relationship between CYP450 genotype and side effects, a longitudinal model including weekly ADR scores was used. To account for the clustering of individuals around recruitment centres, mixed models were used (fitted with maximum likelihood); centre was entered as a random effect along with individual. Fixed effects included age, sex, cytochrome

P450-inhibiting comedication, current depression severity, baseline ADR rating and linear and quadratic effects of time.

The inclusion of baseline depression severity was to ensure that reports of ADRs were not confounded by depressive symptoms (Uher *et al*, 2009a), whilst the inclusion of baseline ADR ratings was necessary to ensure ADRs were in fact associated with the antidepressant being taken, in the context of high endorsement of ADRs prior to treatment (see Chapter 3).

Associations with total ADR burden were tested using linear models, whilst the weekly presence/absence of each specific ADR was examined using logistic models. When considering the specific ADRs, in order to correct for the 21 different outcomes, a Bonferroni correction for multiple hypothesis testing was applied; only associations where $p<0.002381$ were considered significant (0.05/21 = 0.002381).

5.2.6.2 Serum concentration of antidepressant predicting side effects

A longitudinal model was also used to assess the impact of serum concentration of antidepressant on side effects. Four standardised serum measures were considered; concentrations of the antidepressant, its primary metabolite, the metabolite to drug ratio and finally the total concentration of drug plus metabolite. Covariates were as described above, when considering CYP450 genotype prediction of side effects, with the exception that cytochrome P450-inhibiting comedication was not included in the model.

For any ADRs which showed significant associations with serum concentration of antidepressant, secondary analyses were undertaken including dose as a covariate, to probe the nature of the serum-side effect association.

5.2.6.3 Drug metabolising enzymes predicting study drop out

Finally, CYP450 genotype was considered as a predictor of time to study discontinuation, using a survival Cox proportional hazards model. Covariates of age, sex, centre, baseline depression and baseline total ADR score were included in the model.

5.3 Results

5.3.1 Descriptive statistics

The CYP450 genotype frequencies, serum concentration of antidepressants and relationship with dose are described in Chapter 4. Patterns of study drop out are described in Chapter 2. With regards to side effects, it was observed that frequencies differed between the two drugs (Uher *et al*, 2009a). The pattern of overall side effect burden is shown in Figure 5-2, with specific side effects details in Figure 5-3, Figure 5-4 and Figure 5-5.

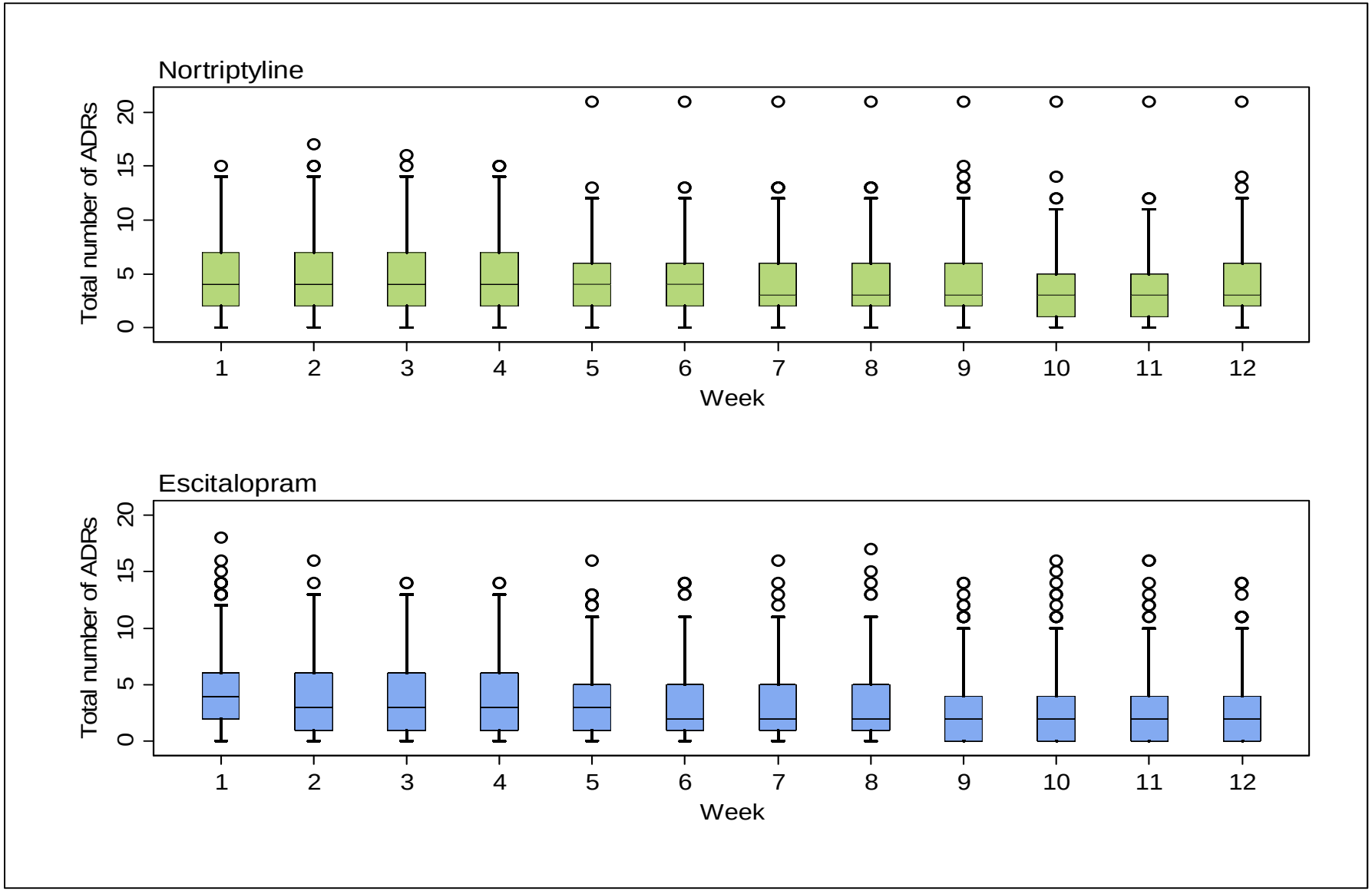

Figure 5-2: Total number of ADRs reported per week, by drug

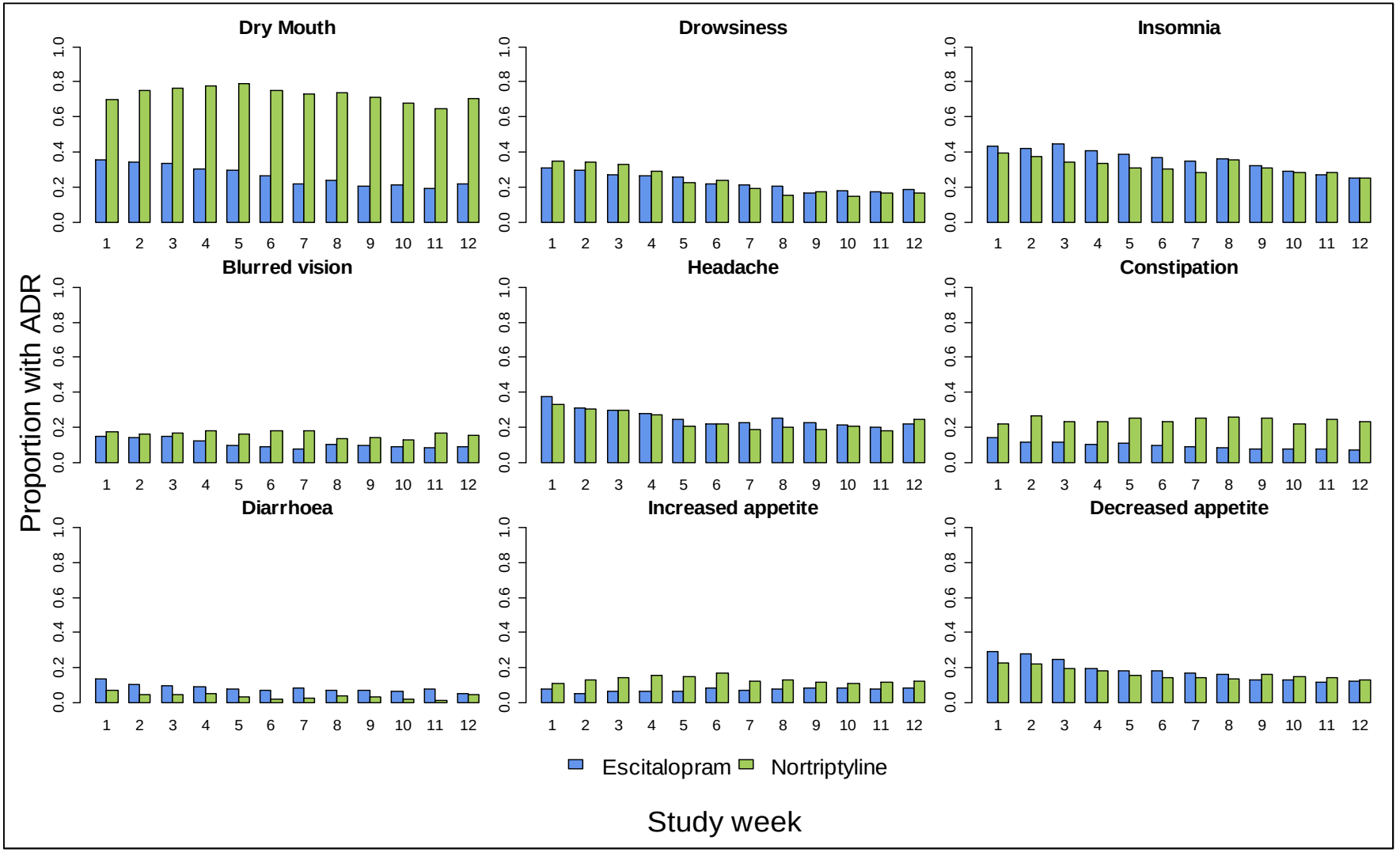

Figure 5-3: Prevalence of specific ADRs per week, per drug (A)

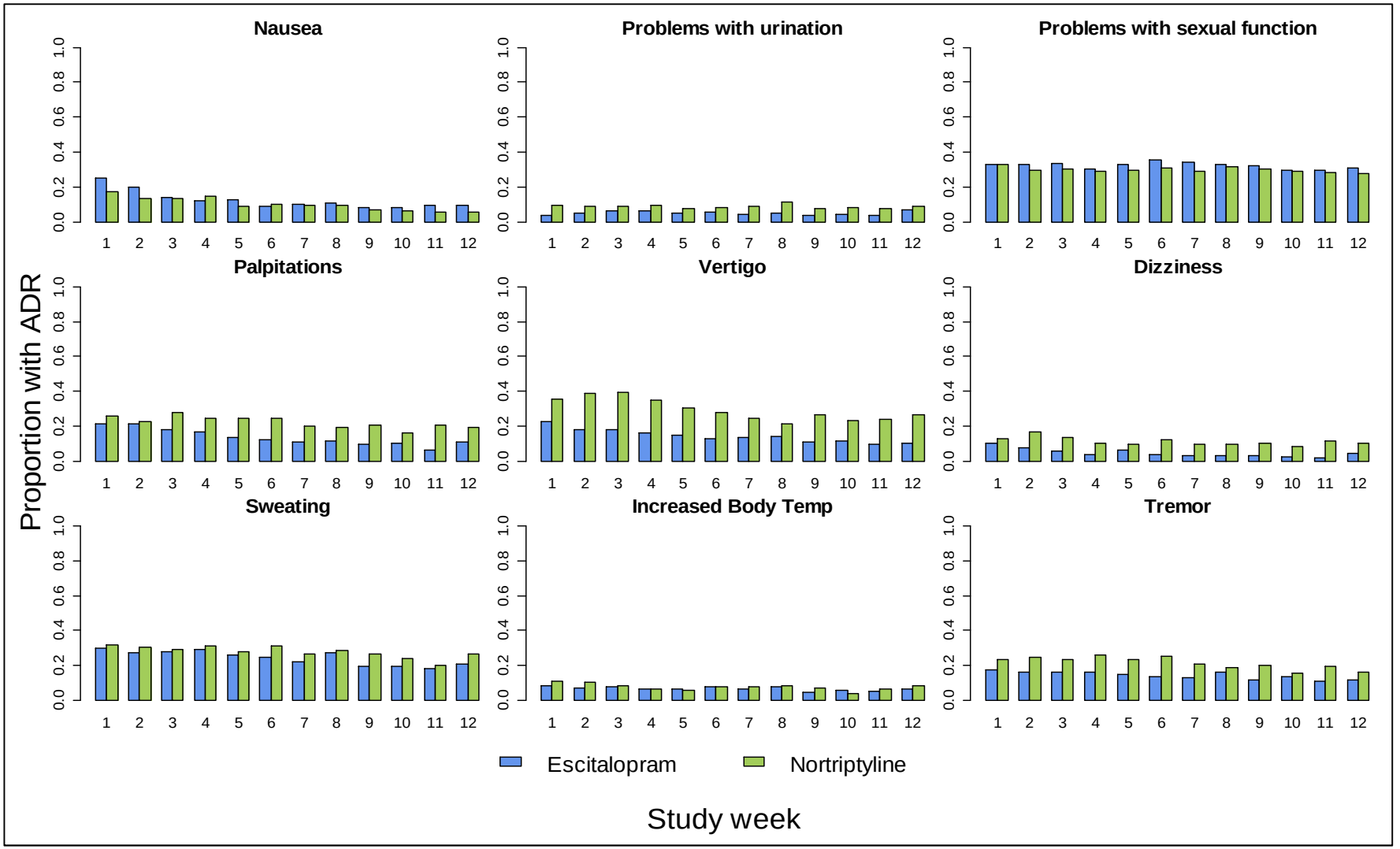

Figure 5-4: Prevalence of specific ADRs per week, per drug (B)

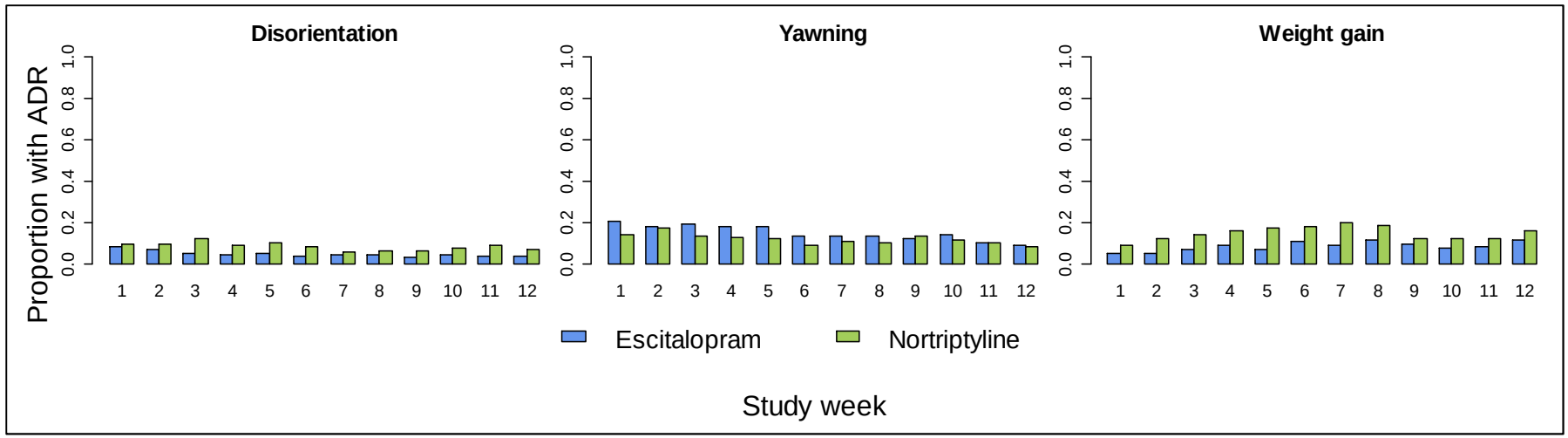

Figure 5-5: Prevalence of specific ADRs per week, per drug (C)

5.4.1.1 Comparison of side effects in patients with and without serum measurements available

Comparisons of baseline measures between patients with and without serum measurements available are also described in Chapter 4. When considering differences in terms of side effects (where measurements were available), amongst patients taking escitalopram, total ADR burden in week eight (when samples were taken for serum measurements) was unrelated to whether serum measurements of antidepressant concentration were available or not. However, total ADR burden was associated with whether serum measurements were available amongst patients taking nortriptyline ($t(236)=-2.13$, $p=0.005$); higher ADR burden was seen in those patients without serum measurements ($M=4.91$, $SD=3.57$, compared with patients with serum measurements; $Mean=3.80$, $SD=3.28$).

5.4.2 CYP450 genotype and side effects

CYP450 genotype predicted neither overall side effect burden nor specific side effects reports, for patients taking either escitalopram or nortriptyline.

5.4.3 Serum concentration of antidepressant and side effects

Furthermore, overall side effect burden was not predicted by any of the four measures of antidepressant serum concentration considered, for patients on either escitalopram or nortriptyline.

Table 5-1 shows the results for prediction of overall side effect burden by both CYP450 genotype and serum concentration of antidepressant.

Table 5-1: Prediction of overall side effect burden; all associations not significant

Antidepressant	Predictor	n	p	coeff	SE
Nortriptyline	*CYP2D6* genotype	251	0.5638	-0.133	0.229
	nortriptyline	184	0.8534	-0.029	0.154
	10-hydroxynortriptyline	180	0.7323	0.063	0.183
	ratio (OH-nortriptyline: nortriptyline)	178	0.1895	0.232	0.176
	total (nortriptyline+OH-nortriptyline)	178	0.7975	0.044	0.170
Escitalopram	*CYP2C19* genotype	340	0.9627	-0.004	0.084
	escitalopram	275	0.0852	0.250	0.145
	desmethylcitalopram	205	0.1587	0.248	0.175
	ratio (desmethylcitalopram:citalopram)	204	0.6945	-0.066	0.167
	total (citalopram+desmethylcitalopram)	204	0.1155	0.260	0.165

For the majority of the 21 specific side effects considered, there was also no significant relationship with serum concentrations of antidepressant (details of non-significant associations for each individual ADR are in the Appendix E). However, there were three exceptions to this.

Firstly, for all patients (taking nortriptyline or escitalopram), a significant association was seen between dry mouth and serum concentrations of both drug, metabolite and total level of drug plus metabolite (Table 5-2), although not ratio of metabolite to drug. When the effect of dose was removed, dry mouth was significantly associated with metabolite concentrations in patients taking nortriptyline (n=164, p=0.0011, OR=1.841, SE=0.345) and total drug plus metabolite concentration (n=163, p=0.0016, OR=2.082, SE=0.484). All other associations were no longer significant.

Table 5-2: Relationship between serum concentrations and dry mouth (significance threshold p<0.00238)

Antidepressant	Serum concentration	n	p	OR	SE
Nortriptyline	nortriptyline	184	0.0023	1.826	0.362
	OH-nortriptyline	180	1.20×10^{-4}	2.100	0.405
	ratio (OH-nortriptyline: nortriptyline)	178	0.0406	1.406	0.234
	total (nortriptyline+OH-nortriptyline)	178	4.97×10^{-5}	2.284	0.465
Escitalopram	escitalopram	275	6.69×10^{-4}	1.480	0.170
	desmethylescitalopram	205	0.0018	1.420	0.159
	ratio (desmethylcitalopram:citalopram)	204	0.6162	0.931	0.133
	total (citalopram+desmethylcitalopram)	204	0.0012	1.496	0.186

Secondly, amongst patients taking escitalopram, significant associations were observed between diarrhoea and the ratio of metabolite to drug (n=203, $p=4.96 \times 10^{-4}$, OR=0.597,

SE=0.088), which remained significant when covarying for dose (n=188, *p*=0.0013, OR=0.632, SE=0.090).

Thirdly, for patients taking escitalopram, dizziness was significantly associated with concentration of metabolite (n=202, p=3.28x10^{-5}, OR=1.564, SE=0.168). This association also remained significant after covarying for dose (n=188, *p*=1.05x10^{-6}, OR=1.658, SE=0.172).

5.4.4 CYP450 genotype predicting study drop out

CYP450 genotype was unrelated to study discontinuation, for patients taking either escitalopram (n=376, p=0.118, Hazard Ratio=0.870, 95% CI=0.731-1.034), or nortriptyline (n=284, p=0.174, Hazard Ratio=1.300, 95% CI=0.891-1.898). The (non-significant) hazard ratios given describe the increase in relative risk of discontinuation, moving from less to more active CYP450 genotypes.

5.5 Discussion

5.5.1 Summary of results

Cytochrome P450 genotype did not predict overall side effect burden, any of the 21 specific ADRs measured, or study discontinuation in this sample. Further investigation, using serum concentration measures indicated that for both overall burden, and the majority of specific ADRs, there was also no relationship with circulating levels of antidepressant. Exceptions to this general pattern were observed for dry mouth, dizziness and diarrhoea.

In the case of dry mouth, higher serum concentrations of antidepressant were linked to higher risk of the side effect for both escitalopram and nortriptyline. After adjusting for the influence of dose on serum concentration, the association with both metabolite and total serum concentration remained significant for patients taking nortriptyline.

Amongst patients taking escitalopram, associations of diarrhoea with the ratio of metabolite to drug, and dizziness with concentration of metabolite were observed; both of these associations also remain significant after adjusting serum concentrations for prescribed dose of escitalopram.

5.5.2 Comparison to previous literature

Whilst the existing literature on the association between drug metabolism rates and ADRs is limited (predominantly relying on smaller samples or case study reports), there are three key papers exploring the issue in well-powered samples.

The publications on the STAR*D sample represent the largest sample to date in which CYP450 genotype has been explored in relation to antidepressant ADRs (n=1,953; Mrazek *et al*, 2011; Peters *et al*, 2008). Both papers used measures of overall tolerance to citalopram treatment using study exit data. This could be compared to either the measures of weekly overall side effect burden, or study drop out within the GENDEP sample. But using either of these

outcomes, no association is observed in GENDEP between CYP450 genotype and ADRs (for either escitalopram or nortriptyline), in line with the report from Peters et al. (2008).

The Rotterdam Study (Bijl *et al*, 2008) also included a large sample size (n=1,198), a range of antidepressants were considered and categorised as either tricyclic antidepressants or SSRIs. Amongst patients taking SSRIs, no association was observed between *CYP2D6* genotype and clinical outcomes. Amongst patients taking tricyclic antidepressants, no association was seen with treatment discontinuation. However, the findings in the Rotterdam Study contradict those presented here, in that there was increased risk of antidepressant switching amongst patients taking tricyclic antidepressants with poor metaboliser genotypes.

In GENDEP it is not possible to study antidepressant switching (the data are not available), in order to examine how robust this disassociation between antidepressant switching and treatment discontinuation is. Nevertheless, by using the detailed measurement of side effects available in GENDEP, together with study drop out data, the work presented here provides an assessment of a range of methods to assess adverse treatment events. Given the range of different ADRs, their varying pharmacological basis, as well as the changing profile of risk to ADRs across treatment (Uher *et al*, 2009a) and their impact of treatment adherence (Mitchell, 2006), this more in depth look in particularly valuable.

5.5.3 Methodological considerations

Many of the methodological considerations for these analyses are similar those also detailed in Chapter 4.

5.5.3.1 Power calculation

Despite finding limited evidence that individual variation in drug metabolism rates is linked to ADRs, power calculations suggest that the analyses presented here are sufficiently powered to detect any effects of a magnitude that would be clinically useful. For example, when considering

CYP450 genotype prediction of overall side effect burden, we used G*Power (Faul *et al*, 2007) to estimate that a sample of 250 patients can detect effect sizes explaining 3.1% of the variance in total ADR burden with 80% power (at a threshold of $p<0.05$); this corresponds to a change of approximately 0.34 points on the ASEC scale.

Nevertheless, the potential translational impact should also be considered in addition to the statistical power available to detect significant associations. Clinical utility of prediction is determined in part by the incidence rate of the outcome. In the case of the three specific ADRs where associations with serum concentrations were observed, both diarrhoea and dizziness are rare side effects (with prevalence below 10% of reports from patients taking escitalopram in this sample). This contrasts with dry mouth, which occurs in more than 70% of reports from patients taking nortriptyline. Therefore the association between serum concentration of antidepressant and dry mouth is likely to be of greater clinical value in terms of the number of patients affected.

5.5.3.2 Multiple-hypothesis testing

Whilst Bonferroni corrections were applied to consider the 21 specific ADRs tested, no further corrections were applied for the number of predictors tested (as these variables are interdependent). Thus, the threshold of significance applied within these analyses was lenient. Despite this, there was still limited evidence of a link between drug metabolism rates and antidepressant side effects.

5.5.3.3 CYP450 genotypic categorisation

As in Chapter 4, alternative methods of the categorisation of CYP450 genotypes were explored. Four or two category models for *CYP2C19* (Brandl *et al*, 2014; Peters *et al*, 2008) and two category models for *CYP2D6* (Mrazek *et al*, 2011) were considered. The results remained unchanged; CYP450 genotype was not predictive of overall ADR burden, specific ADRs or study drop out for either drug.

5.5.4 Study limitations

5.5.4.1 Biases arising from missing data

As mentioned in Chapter 4, if missing data (either due to study drop out, or not providing blood samples for serum measurements) is not at random, this may be a potential source of bias in the study. If data is more likely to be missing from patients who experienced high levels of ADRs, then it made lead to an underestimation of the impact of drug metabolism variability on side effects. However, the analysis of study drop out goes some way to addressing this issue. Given samples were taken at baseline to measure CYP450 genotype, and study drop out data is available for all participants, it is possible to accurately assess whether CYP450 genotype is linked to study drop out; we find no evidence to support this hypothesis. However, this does not exclude the possibility of missing data being linked to other variables of interest.

5.5.4.2 Flexible dosing protocol

Patients received treatment according to a flexible dosing protocol, where both treatment response and side effects could be used by clinicians to inform dose alterations throughout the study. Therefore the doses received by patients were adjusted in response to treatment outcomes as the study proceeded. This means that the findings presented here indicate the extent to which data on CYP450 genotypes and serum concentrations of antidepressants predict ADRs when used in addition to clinical observation.

5.5.5 Conclusions

Where antidepressant dosage is monitored by clinicians and adjusted using their judgement during the first eight weeks of treatment, CYP450 genotypes do not explain treatment associated side effects, as measured by overall side effect burden, specific ADRs, or study discontinuation. However there is some evidence that serum concentration of antidepressants may be linked to the occurrence of dizziness and diarrhoea for patients taking escitalopram and dry mouth for patients taking nortriptyline.

Chapter 6 Transcriptomics and the mechanisms of antidepressant efficacy

6.1 Introduction

6.1.1 Molecular mechanisms of antidepressant action

With antidepressant medications, the blockade of neurotransporters occurs immediately after drug administration. But, as discussed in Chapter 1, there is often a significant delay of 2-3 weeks before symptom improvements can be observed (Frazer and Benmansour, 2002; Uher *et al*, 2011). Therefore, the molecular mechanisms that underlie the therapeutic action of these drugs remain unclear.

It seems that the immediate action of antidepressants within the synapse activates a pathway of downstream adaptive changes necessary for clinical improvement; the precise nature of these changes is unclear but they may occur through transcription pathways altering gene expression levels (Duman *et al*, 1997; Lesch *et al*, 2002).

6.1.2 Potential involvement of gene expression

Using candidate gene approaches, gene expression alterations that are associated with treatment response have been observed within several key systems linked to antidepressant action; including inflammatory (Powell *et al*, 2013), stress response (Cattaneo *et al*, 2013) and neuroplasticity (Belzeaux *et al*, 2010; Cattaneo *et al*, 2010) pathways (see Introduction for further details).

By focusing specifically on those changes in gene expression that are associated with treatment outcomes, these studies are able to probe the mechanism by which these drugs exert their therapeutic effects (Gerhold *et al*, 2002). This is particularly important in light of the high degree of variability seen between patients in terms of efficacy (Trivedi *et al*, 2006).

6.1.3 Transcriptome-wide approaches

However, the majority of patient studies have used candidate gene approaches to measuring expression levels. By using microarray technology, it is possible to interrogate the entire

transcriptome in a systematic and hypothesis-free manner. This enables the detection of novel gene expression changes which are associated with antidepressant efficacy, and places evidence regarding changes in expression levels for candidate genes within the context of the whole transcriptome (albeit it at the cost of a higher multiple-hypothesis testing burden, necessitating larger sample sizes).

Indeed transcriptome-wide analyses using animal models indicate that gene expression changes associated with antidepressant action are likely to be complex, involving the co-ordinated change of many different transcripts to produce therapeutic effects (Sillaber *et al*, 2008).

Nevertheless, to date there is limited work exploring expression changes linked to treatment response in patients on a transcriptomic level. Mamdani et al. (2011) used blood samples from 63 patients with major depressive disorder to identify 32 probesets that were differentially expressed according to response to citalopram, the most significant of which was Interferon Regulatory Factor 7 (*IRF7*). Only two other studies have so far been published looking at transcriptome-wide gene expression changes in patients receiving antidepressant treatment, and these had very sample sizes (less than 10 patients; Belzeaux *et al*, 2012; Kálmán *et al*, 2005).

Transcriptome-wide information not only allows individual, gene-by-gene analysis, but can also be used within a system-based approach, where networks of genes which share expression patterns can be investigated (Langfelder *et al*, 2008). It is known that genes show structured correlation in expression levels (Lee *et al*, 2004); by using coexpression based approaches it is possible to organise genes into co-regulated modules which are likely to be functionally related and identify alterations in expression at a network level. This has the added benefit of reducing the multiple hypothesis testing burden of a gene-by-gene approach and so increases statistical power to detect effects.

This network-level methodology has not yet been applied to gene expression changes in patients receiving antidepressant treatment.

6.1.4 Aims

In this chapter, the antidepressant effects on gene expression were investigated using both an individual gene and a network-based approach. Variability in gene expression changes were considered as potential correlates of antidepressant therapeutic effect.

Given the divergent mechanisms of action of the two antidepressants used in GENDEP, together with the evidence from animal work that expression profiles may vary between different treatments (Palotás *et al*, 2004), drug-specific analyses were also performed. This allowed exploration of whether any significant associations between gene expression changes and treatment outcomes were common across antidepressant drug classes, or were drug specific.

6.2 Methods

6.2.1 Participants

Participants included in this chapter are drawn from the GENDEP sample, as described in detail in Chapter 2.2. Figure 6-1 shows the patients included in analyses presented in this chapter.

6.2.2 Treatment response measures

Treatment response was treated as a quantitative trait using percentage change in MADRS from week zero to week eight (when blood samples were taken), adjusting for age and centre (see Chapter 2.4.1 for further details).

At week zero, MADRS scores amongst the subsample of GENDEP included in this analysis were Mean=29.28, SD=6.44, whilst at week eight of treatment, MADRS scores were Mean=15.22, SD=9.12. There was no significant association between baseline scores and adjusted percentage change in MADRS.

6.2.3 Transcriptome-wide gene expression data

Methods of sample processing and transcriptomic quality control steps are described fully in Chapter 2.7. Briefly, whole blood samples were obtained at week zero and week eight of treatment for 136 patients. After standard RNA extraction procedures, samples were analysed using HumanHT-12 v4 Expression BeadChip microarrays (Illumina, Inc., San Diego). After quality control measures were applied to the data, a total of 242 paired samples, from 121 individuals remained, with gene expression measurements from 29,765 probes.

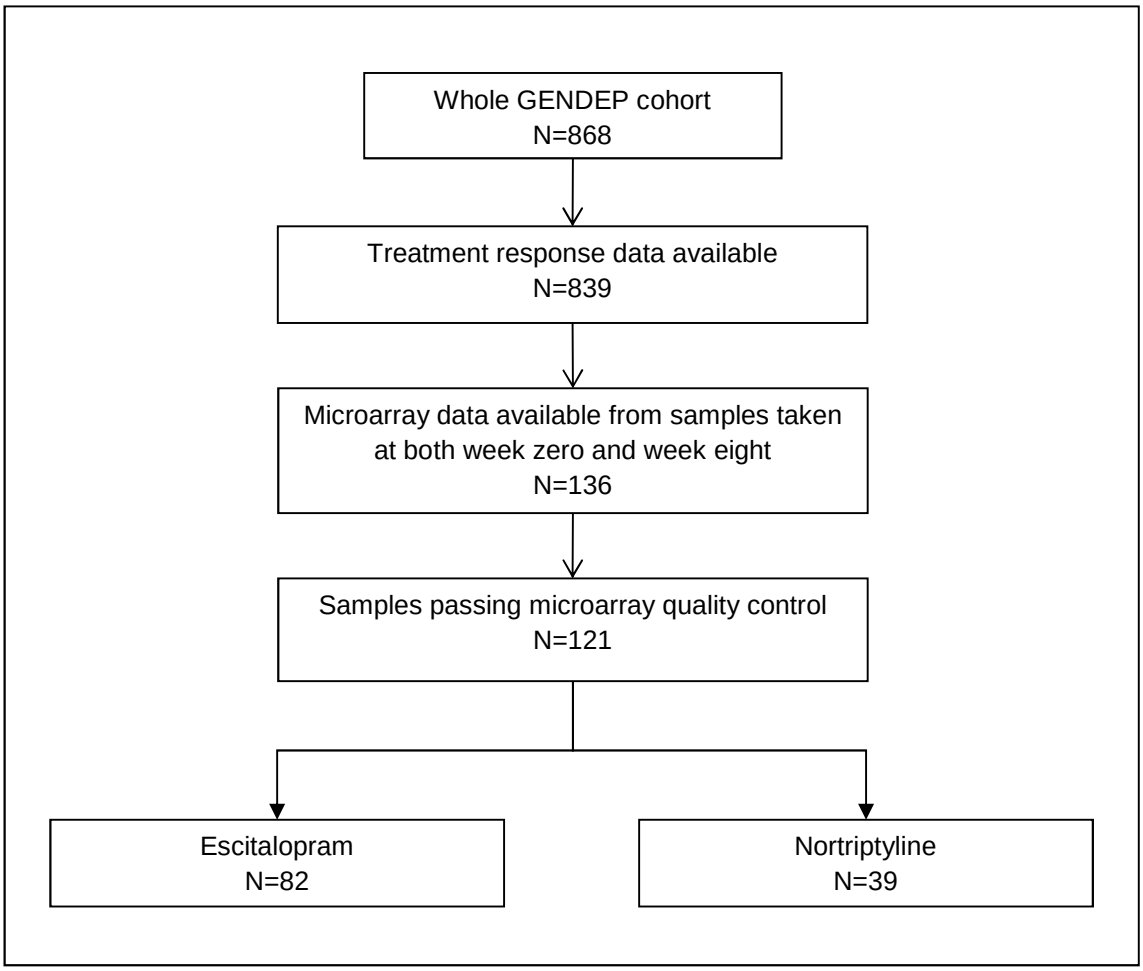

Figure 6-1: Sample included in Chapter 6

6.2.3.1 Cell type proportions

As cell counts for blood samples were unavailable, a deconvolution method was used to estimate cell proportions in each sample. Using the R package CellMix (Gaujoux *et al*, 2013), previously identified cell-type specific gene expression profiles (Abbas *et al*, 2009) were used to estimate the proportion of neutrophils, monocytes and lymphocytes in each sample. Paired t-tests were used to compare the estimated proportions of in samples taken at week zero compared to week eight. No significant differences were observed for the three cell types considered; lymphocytes t(120)=-0.6843, p=0.4951, neutrophils t(120)=1.3312, p=0.1856, or monocytes t(120)= -0.9833, p=0.3274.

Additionally, there was no significant correlation between treatment response (as measured using adjusted percentage change in MADRS score from baseline to week 8) and change in cell

proportions; lymphocytes, r(199)=-0.0960, *p*=0.2951, neutrophils, r(199)=0.0177, *p*=0.8469, monocytes, r(199)=0.0902, *p*=0.3250.

6.2.4 Statistical Analysis

In order to examine the association between gene expression changes and response to antidepressant treatment, two complementary approaches were used; first analysing individual genes and then analysing networks of coexpressed genes.

6.2.4.1 Analysis of individual genes

In the first approach, each measured gene expression level was considered independently as a potential correlate of therapeutic response. Changes in gene expression for each gene probe were correlated with antidepressant treatment response. We present the false discovery rates (FDR), as calculated within the *limma* package (Smyth, 2005), using the method of Benjamini and Hochberg (1995). To correct for multiple hypothesis testing, we considered only those associations with FDR (or q-value) <0.05 as significant, however all associations with $p<0.001$ have been reported as suggestive.

In order to annotate the functions of the genes associated with antidepressant treatment and response, additional pathway analysis was undertaken using IPA (Ingenuity® Systems), considering all gene probes where $p<0.01$. The IPA network displaying the closest overlap with the inputted gene probes is described, along with the IPA network score. The IPA network score is based on the –log (Fishers exact test) measuring the fit between the molecules identified within the data and the IPA network. Therefore higher IPA network scores indicate closer overlap between the IPA network and the inputted gene probes; see http://www.ingenuity.com/ for further details. Brief details on the IPA-defined functional categories referenced within this chapter are included in Appendix F.

6.2.4.2 Analysis of networks of coexpressed genes

In the second methodological approach, a weighted co-expression approach (WGCNA; Langfelder *et al*, 2008) was applied, in order to identify coexpression modules (that is, clusters of gene probes with highly correlated gene expression levels).

Modules of coexpression were built using the gene expression profiles obtained at week zero. Using the correlation strength of each gene probe to all other gene probes (as measured by the connectivity measure *k*), the top 10,000 most connected gene probes within this dataset were selected. This matrix was then transformed into a weighted adjacency matrix using soft thresholding (raising the absolute values of the correlation matrix to the power β).

The topological overlap for all pairs of gene probes was then calculated (which describes the degree of dissimilarity between two gene probes in terms of their patterns of connectivity with all other gene probes), and hierarchical clustering was performed on the 1-topological overlap matrix. Each module was defined as a branch of this hierarchical clustering tree.

To summarise the gene expression profile within each identified module, the module eigengene (or first principal component) was calculated for the gene probes belonging to the module. Using the modules identified in the week zero gene expression data, the module eigengene was calculated for each module in both the week zero and the week eight gene expression data. From this, the change in module eigengene (or change in gene expression profile during treatment) was obtained for each module, for each individual. The change in module eigengene for each individual was then correlated with antidepressant treatment response. To account for the number of modules tested, a threshold for significance of FDR<0.05 was used.

In the modules showing the most significant correlation with treatment response, module membership and gene significance values were calculated for each gene probe included. In this dataset, we defined module membership as the correlation between gene expression and the

module eigengene at week zero, and can be used to identify those gene probes with greater connectivity within the module. Gene significance describes the correlation between the change in gene expression and treatment response.

We explored the relationship between module membership and gene significance, then used these measures to define those gene probes within the top 50% for both gene significance and module membership as “hub genes”.

These hub genes were used to annotate the modules of interest, using IPA (Ingenuity® Systems). The IPA networks displaying the strongest overlap with the inputted gene probes are described, along with the associated network scores (where higher scores indicate stronger overlap).

6.2.4.3 Drug specific analyses

The primary analysis used the entire sample available. However, the two antidepressants used in GENDEP (escitalopram and nortriptyline) have divergent mechanisms of action, and so we also conducted exploratory drug-specific analyses, in order to investigate any drug-specific effects on gene expression.

6.3 Results

6.3.1 Individual Gene Analysis

6.3.1.1 Whole sample analysis

No significant correlations were found between change in gene expression and treatment response, although suggestive associations ($p<0.001$) were observed for 58 probes; see Appendix F. All gene probes where correlation with treatment response was $p<0.01$ were entered into Ingenuity pathway analysis. The IPA network "Haematological Disease, Organismal Injury and Abnormalities, Developmental Disorder" showed greatest overlap with these gene probes (IPA score=46).

Table 6-1 shows the molecules identified within this IPA network. Each IPA network contains 35 molecules; the "network eligible molecules" are those that were identified as suggestive in the GENDEP dataset, with "other molecules" being those identified as interacting molecules in the Ingenuity Knowledge Base. As described in the methods, the IPA network score is based on the –log (Fishers exact test) measuring the fit between the molecules identified within the GENDEP data and the IPA network. Therefore higher IPA network scores indicate closer overlap between the IPA network and the inputted gene probes. The "Top Functions" column describes the three most significant functions for each network. The connections between these molecules, as described using Ingenuity, are drawn in Figure 6-2.

To set this network in context, the details of the top five networks identified are included in Appendix F.

Confirmation of treatment adherence through analysis of antidepressant concentrations in plasma samples was available for 81 patients with gene expression data (this represents 67% of the sample analysed in this chapter). When analyses were restricted to these cases, the pattern of results was similar to that in the whole sample analyses; no gene probes reached significance (FDR<0.05), 93% of gene probes with suggestive association in the whole sample

showed nominal significance (p<0.05) in this subset, and all correlations were in the same direction.

Table 6-1: Details of the IPA network identified within the individual gene analysis of the whole sample

Analysis	Molecules in Network			Score	Top diseases and functions
	Network eligible		Other		
Whole sample	28	BSG COX5B COX6A1 DAD1 DMTN EPB42 ERAP2 FECH FKBP8 GFER GNL2 GYPB HBD HERC2 ICAM4 KLF1 MAGEA3/MAGEA6 MPRIP PSMA7 PSMB4 PSMB7 PSMB8 PSMG3 RIOK3 SLC4A1 SNCA TRIM69 VIMP*	20s proteasome hemoglobin Immunoproteasome Pa28/20s NFkB (complex) PFK Proteasome PA700/20s PSMB	46	Hematological Disease, Organismal Injury and Abnormalities, Developmental Disorder

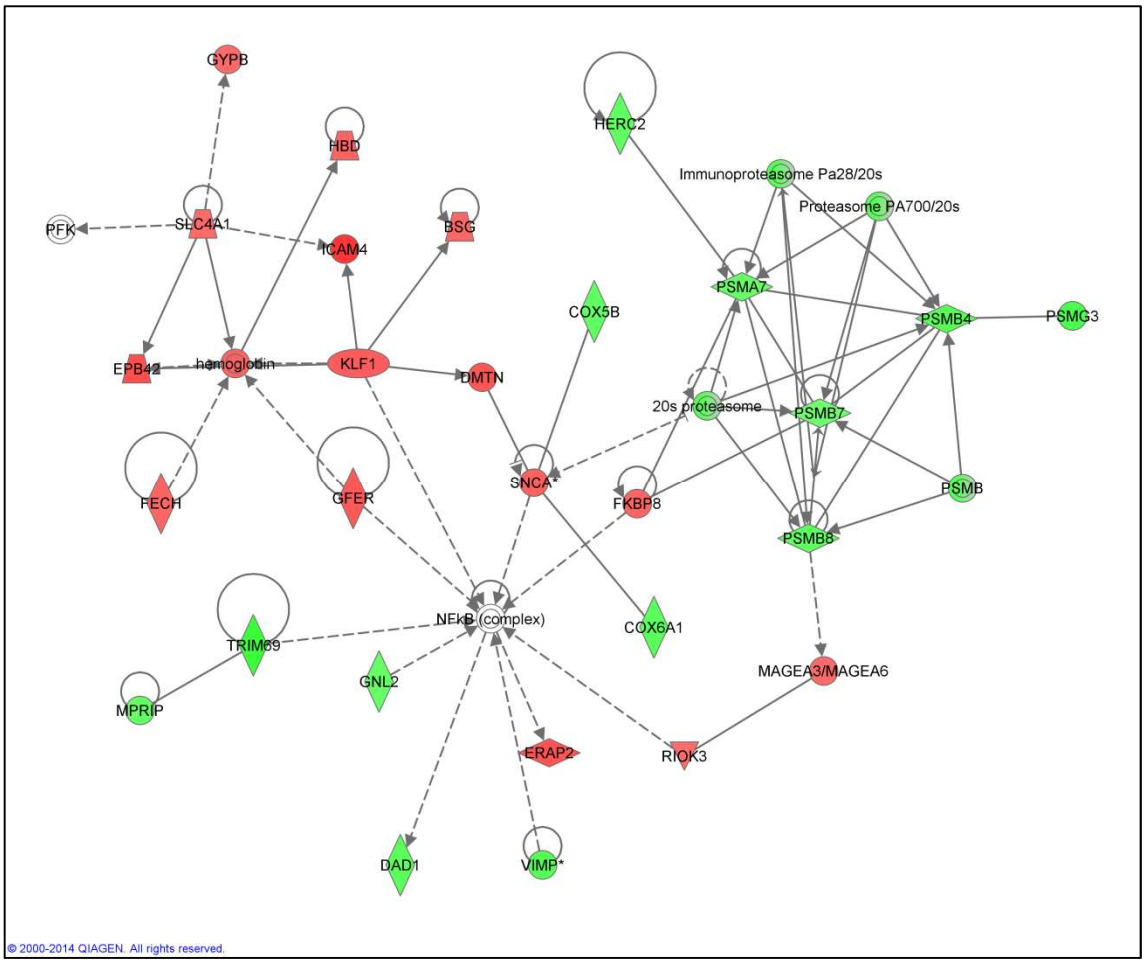

Figure 6-2: Connections between molecules identified in the IPA network identified in the individual gene analysis of the whole sample. Molecules in red show positive correlation with treatment response, whilst molecules in green show negative correlation. Molecules in white are those which are identified as interacting molecules in the Ingenuity Knowledge Base.

6.3.1.2 Escitalopram-specific analysis

For patients on escitalopram (n=82), again no gene probes displayed significant correlations between change in gene expression and treatment response (20 gene probes showed suggestive correlations, see Appendix F). Pathway analysis indicated the network showing greatest enrichment was “Infectious Disease, Organismal Development, Connective Tissue Disorders” (IPA score=37). The molecules within this network are shown in Table 6-2. The connections between these molecules are drawn in Figure 6-3. Brief details of the top five networks identified in the pathway analysis are shown in Appendix F.

Table 6-2: Details of the IPA network identified within the escitalopram-specific sample in the individual gene analysis

Analysis	Molecules in Network			Score	Top diseases and functions
	Network eligible		Other		
Escitalopram	19	AGFG1 AIFM3 ALAS1 CTSB FECH FEN1 GJC1 GRK1 HYAL2 IBSP KMT2E NPM3 NR4A1 PPP2R2B RBP2 RTN4 WISP3 ZC3HAV1 ZFP36L1	ADRB Akt caspase CDO1 Collagen type I Collagen(s) Creb Cyclin A E2f hemoglobin Ikb LDL PDGF BB Pkc(s) PP2A Rxr	37	Infectious Disease, Organismal Development, Connective Tissue Disorders

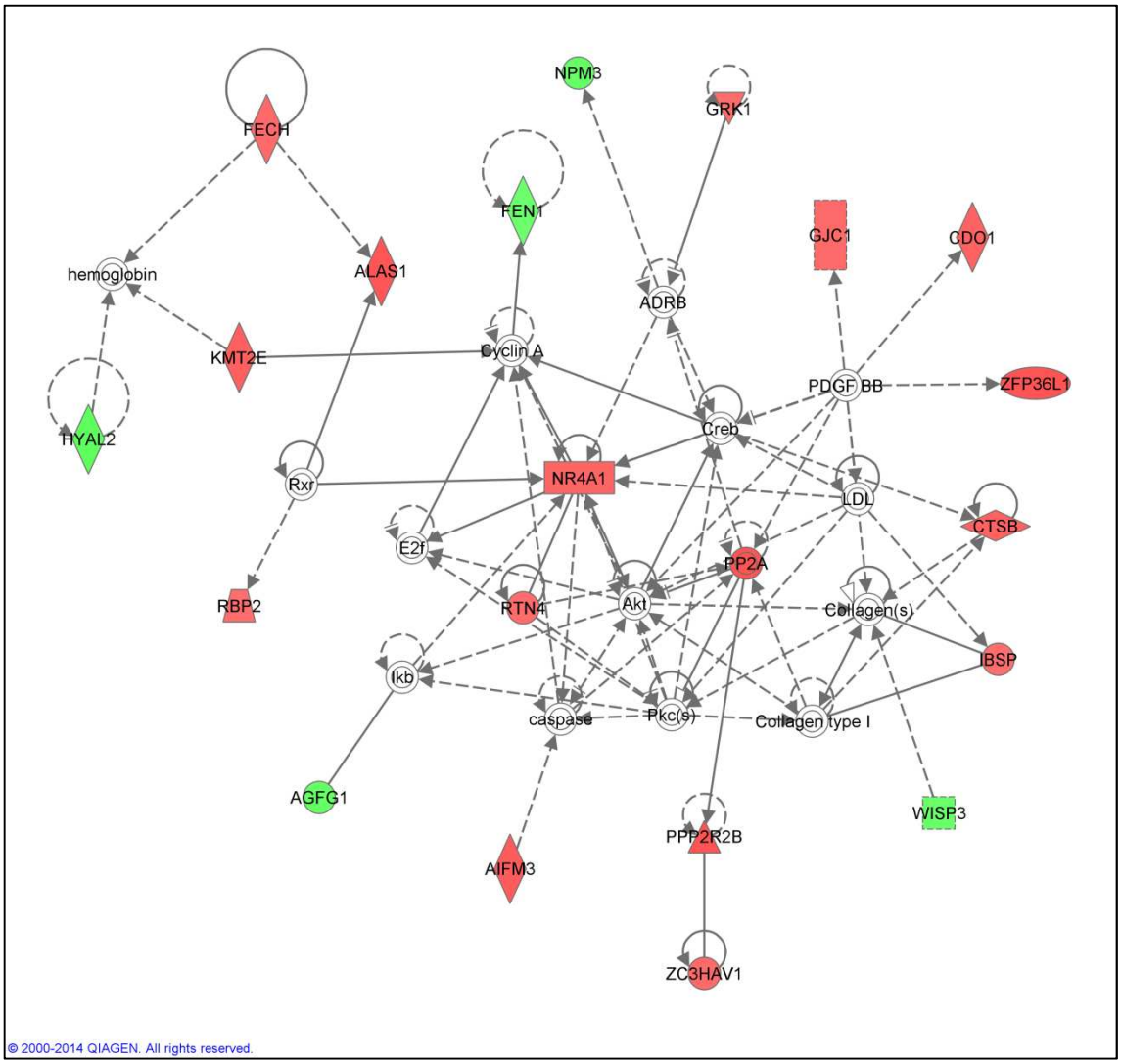

Figure 6-3: Connections between molecules identified in the IPA network identified within the escitalopram-specific sample in the individual gene analysis. Molecules in red show positive correlation with treatment response, whilst molecules in green show negative correlation. Molecules in white are those which are identified as interacting molecules in the Ingenuity Knowledge Base.

6.3.1.3 Nortriptyline-specific analysis

For patients taking nortriptyline (n=39), two gene probes showed gene expression change significantly correlated with treatment response; *MMP28* (probe ILMN_1790951) and *KXD1* (probe ILMN_1790951), as shown in Table 6-3 and Figure 6-4.

Table 6-3: Significant correlations between change in gene expression and treatment response amongst patients taking nortriptyline

Gene	ID	Corr	P Value	Q Value	Chromosome Location
MMP28	ILMN_1752952	-0.683	1.662×10^{-6}	0.031	17q12b
KXD1	ILMN_1790951	0.678	2.103×10^{-6}	0.031	19p13.11c

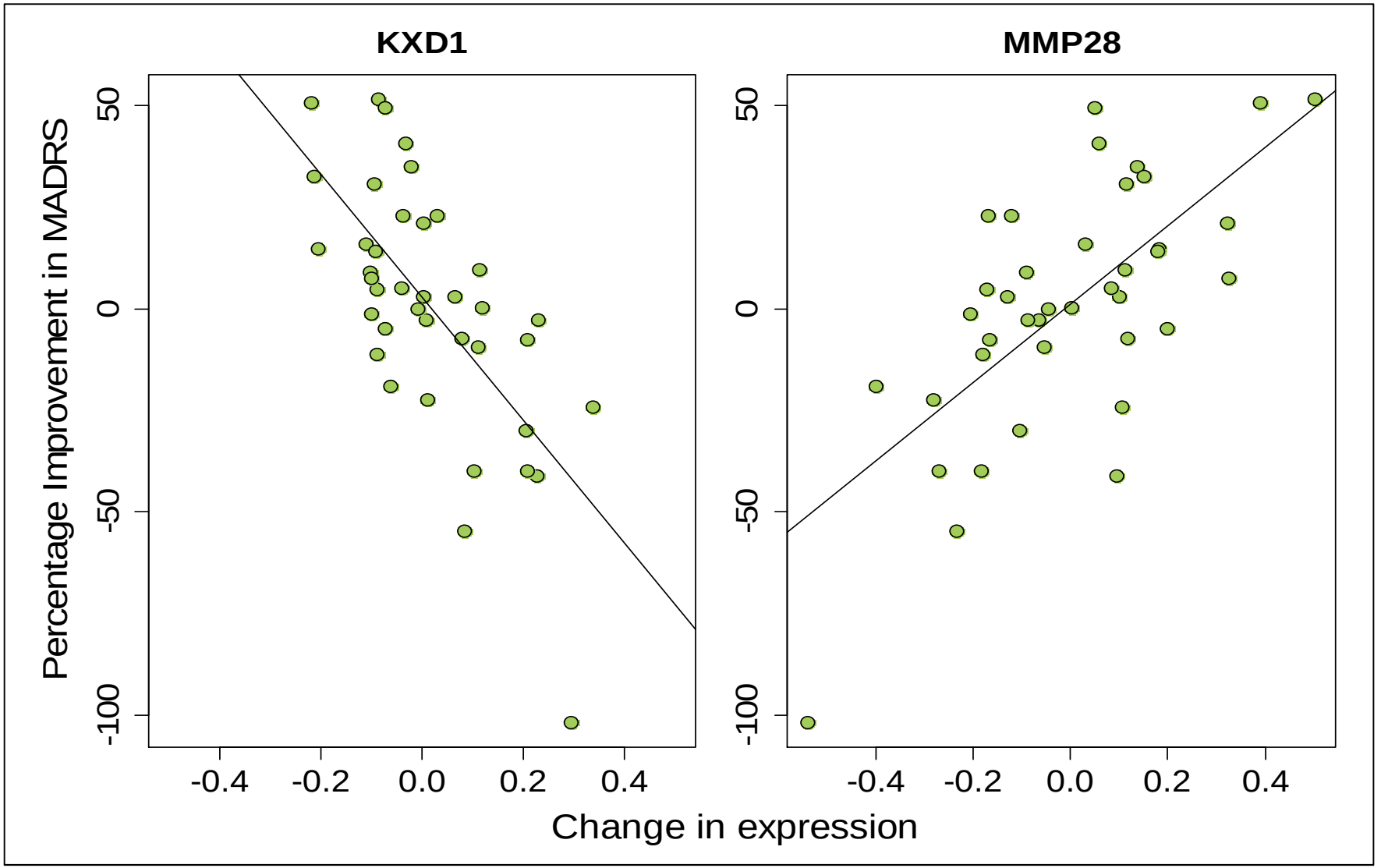

Figure 6-4: Correlation between change in gene expression and percentage improvement in MADRS, for MMP28 and KXD1. (Nortriptyline-specific analysis)

A Fisher's exact test was used to compare overlap between gene probes nominally associated (p<0.05) with nortriptyline or escitalopram, but no significant overlap was observed (p=0.2251).

An additional 59 gene probes showed suggestive associations (see Appendix F), with pathway analysis implicating the IPA network "Cellular Assembly and Organisation, Nervous System Development and Function, Cell Signalling" (IPA score=36). The molecules identified in this network are shown in Table 6-4, with the connections between molecules drawn in Figure 6-5. Brief details of the top five networks identified using IPA are shown in Appendix F.

Table 6-4: Details of the IPA network identified within the nortripytline-specific sample in the individual gene analysis

Analysis	Molecules in Network			Score	Top diseases and functions
	Network eligible		Other		
Nortriptyline	15	AKR1A1 CCNY CDIPT CTDSP1 DDX19A ECHDC1 GMPS MTRF1L NBPF10 (includes others) NDUFAF3 POFUT2 SPTBN2 TRMT44 UBAC1 ZBTB18	AGPAT2 CDCA3 COASY EFHD2 FAM115A FASTKD5 GHITM HAUS6 MTAP NDUFA10 NDUFAF4 NFXL1 POMGNT1 RNF123 TCEAL1 TSTA3 UBC UNC13B USP22 YME1L1	36	Cellular Assembly and Organization, Nervous System Development and Function, Cell Signaling

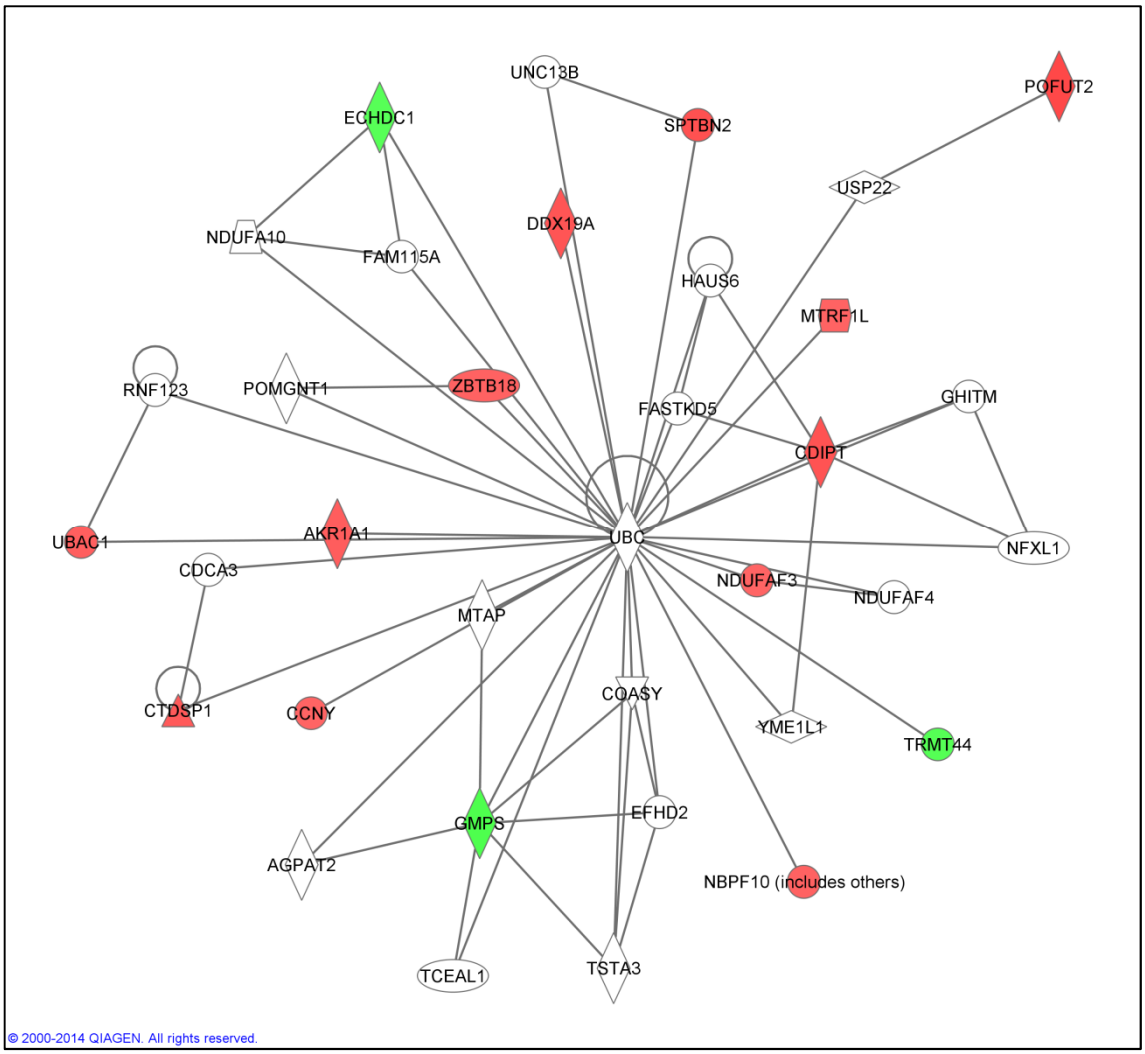

Figure 6-5: Connections between molecules identified in the IPA network identified within the nortripytline-specific sample in the individual gene analysis. Molecules in red show positive correlation with treatment response, whilst molecules in green show negative correlation. Molecules in white are those which are identified as interacting molecules in the Ingenuity Knowledge Base.

6.3.2 Analysis of networks of coexpressed genes

6.3.2.1 Whole sample analysis

Using the gene expression data available at week zero, ten modules of coexpression were identified within the WGCNA analysis framework.

Correlating changes in the module eigengene from week zero to week eight with treatment response, a significant (FDR<0.05) correlation was observed for one module (cor=0.27, p=0.0029, FDR=0.0317), shown in Figure 6-6. This module (referred to as module A) does not show significant correlations with any baseline patient characteristics (see Figure 6-7).

Module A contains a total of 146 genes, of which 141 could be annotated. We observed a substantial and significant correlation (r=0.56, p=1.47x10^{-13}) between module membership and gene significance within module A, showing gene probes with greater connectivity within the module tended to have a stronger association with treatment response (Figure 6-8).

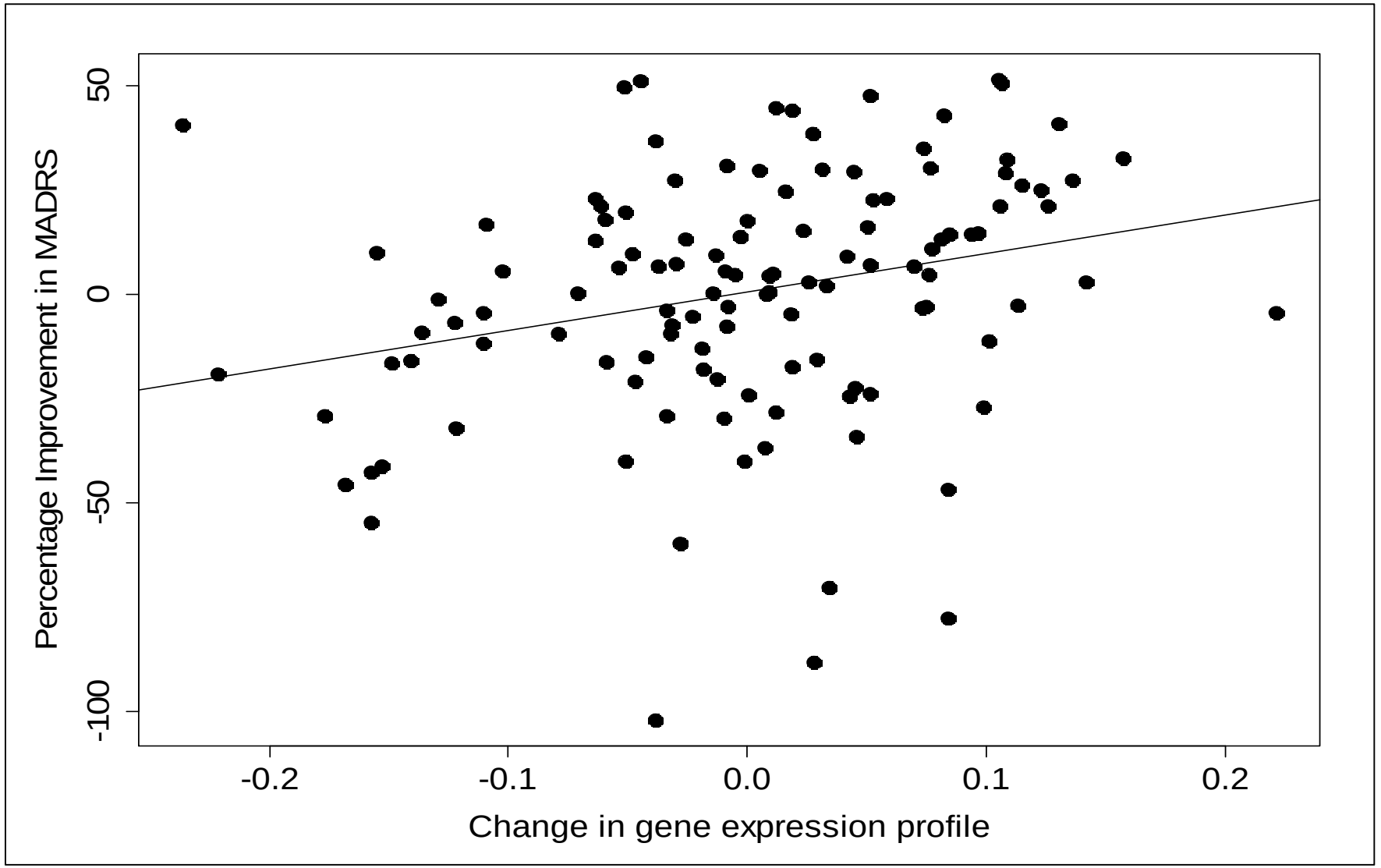

Figure 6-6: Correlation between change in the module eigengene and percentage improvement in MADRS, for module A. (Cor=0.27, p=0.0029)

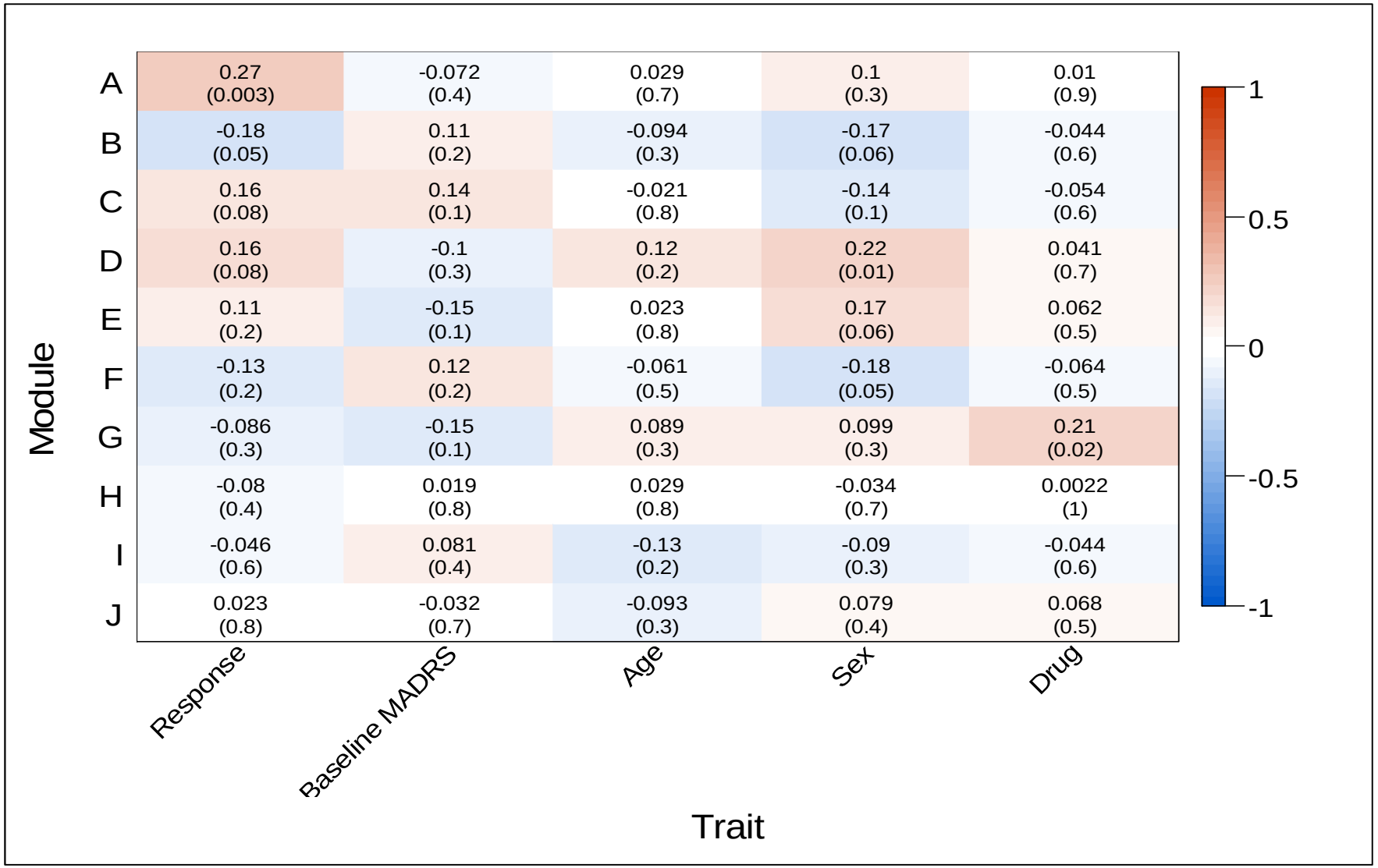

Figure 6-7: Relationship between module eigengenes and traits of interest of each of the ten identified modules.

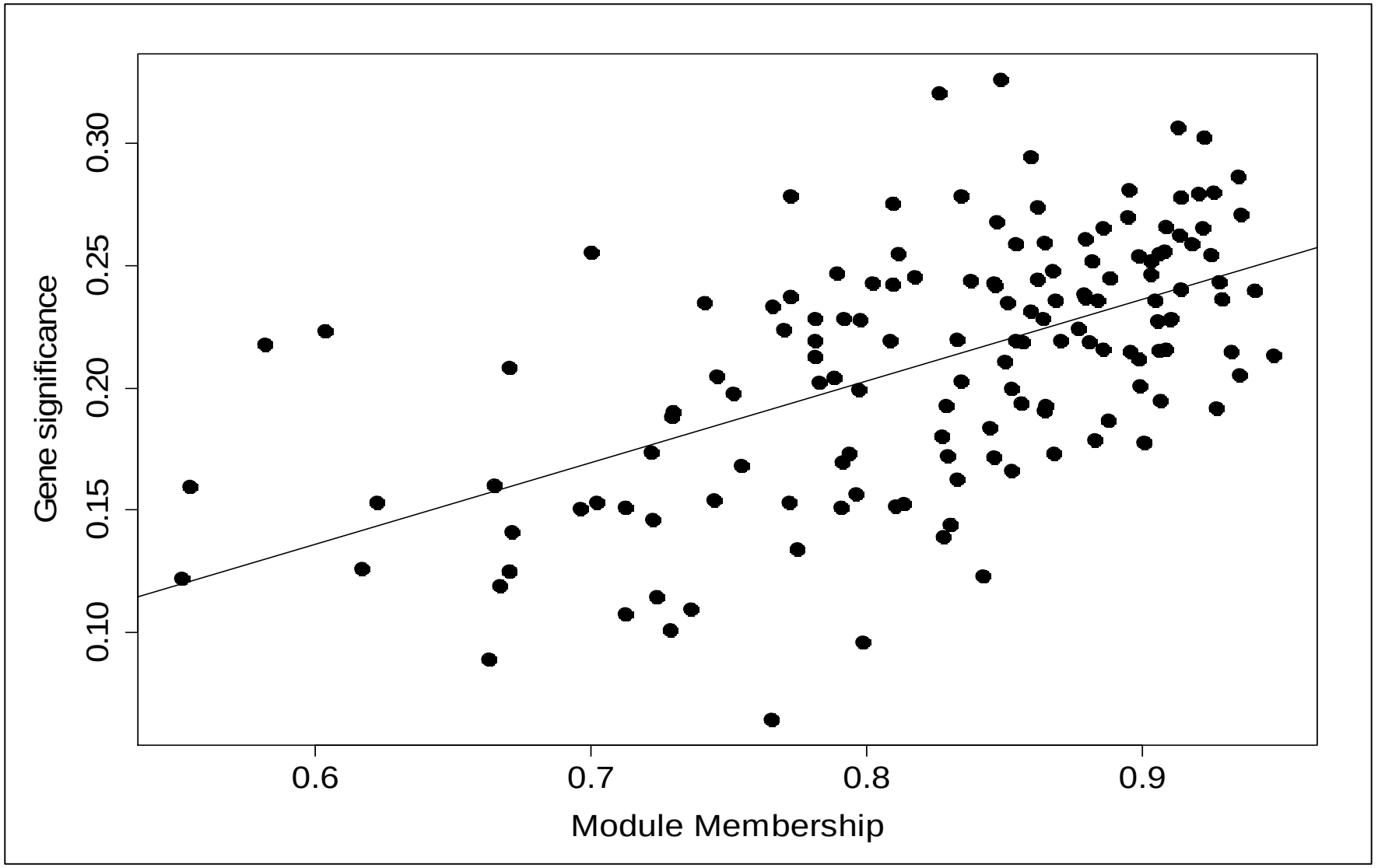

Figure 6-8: Relationship between module membership and gene significance for each gene probe within module A

Pathway analysis using “hub genes” (that is, those gene probes that were within the top 50% for both module membership and gene significance, n=50) revealed the top network for enrichment within module A to be “Cancer, organ development and organismal injury and abnormalities” (IPA score=59). Details of this network are shown in Table 6-5 and Figure 6-9, with further information on the top five networks included in Appendix F. There was no enrichment observed within the module for genes from annotated brain derived modules included in the *BrainLists*, within the WGCNA package.

Table 6-5: Details of the IPA network identified using the hub genes in module A in the whole sample

Analysis	**Molecules in Network**			**Score**	**Top diseases and functions**
	Network eligible		**Other**		
Whole sample	23	ADIPOR1* ASCC2 BCL2L1 CDC34 DMTN EPB42 GMPR GYPB GYPE HBD IFIT1B KLF1 MAP2K3 MPP1 PIP4K2A PLEK2 PSMF1 RIOK3 RNF11 SLC4A1 SNCA* TNS1 YBX3	26s Proteasome Ck2 ERK1/2 Hemoglobin IL12 (complex) Immunoglobulin Interferon alpha NFkB (complex) P38 MAPK PI3K (complex) STAT5a/b Tnf (family)	59	Cancer, Organ Development, Organismal Injury and Abnormalities

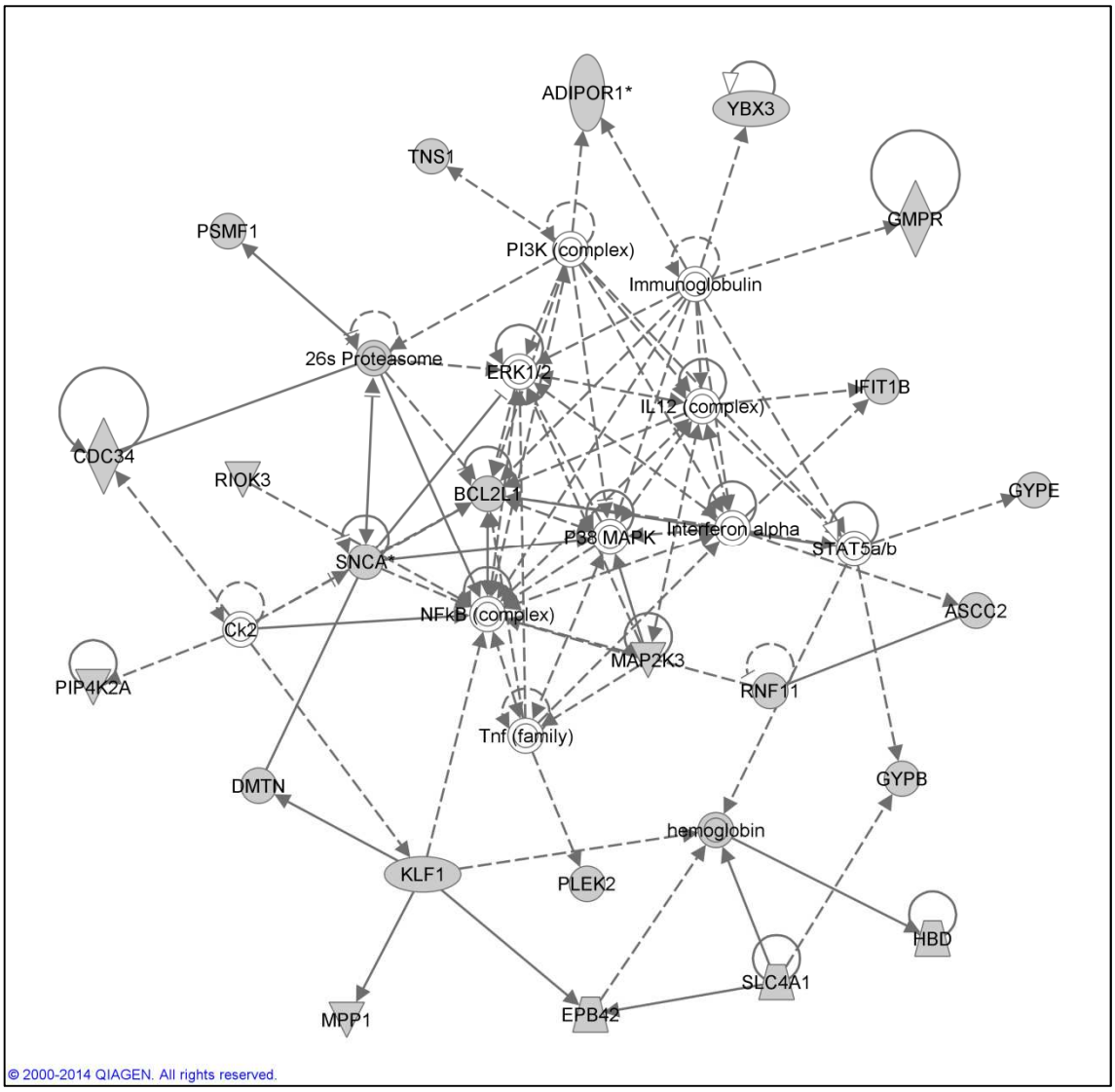

Figure 6-9: Connections between molecules within the IPA network identified using the hub genes in module A in the whole sample. . Molecules in grey are hub genes in the WCGNA module, molecules in white are identified as interacting molecules in the Ingenuity Knowledge Base

6.3.2.2 Escitalopram-specific analysis

Eleven co-expressed modules were identified when considering only those patients who were prescribed escitalopram. No modules showed significant correlation between change in module eigengene and treatment response after correction for the number of modules tested (FDR<0.05). The most significantly associated module (r=0.23, *p*=0.0416, FDR=0.4567) contained 92 gene probes, 90 of which could be annotated. The correlation between module membership and gene significance was r=0.32, *p*=0.0021). Pathway analysis using “hub genes” within this module (n=26) revealed the top network for enrichment to be “Cancer, organ development and organismal injury and abnormalities” (IPA score=37). The details of the molecules identified in this IPA network are shown in Table 6-6 and Figure 6-10, with additional

information on the top 5 networks identified in Appendix F. This is the same network annotation identified in the analysis of the whole sample. Furthermore, of the 50 hub genes identified in the whole sample analysis, 25 were also identified as hub genes in the escitalopram-specific analysis.

Table 6-6: Details of the IPA network identified using the hub genes in most significantly associated module in the escitalopram-specific sample

Analysis	Molecules in Network			Score	Top diseases and functions
	Network eligible		Other		
Escitalopram	14	DCAF12 EPB42 FAM210B GSPT1 IGF2BP2 MARCH8 MKRN1 OSBP2 RBM38* SELENBP1 SLC4A1* STRADB TMOD1 TSPAN5	C4orf32 Ca2+ DCAKD DESI2 ELAVL1 FOPNL GDPGP1 KIAA2013 LRRC37B MARVELD2 OR6T1 RNF212 SLC18B1 SLC41A1 TGFB1 TMEM127 TMEM209 UBC VSIG10 ZNF529 ZSWIM1	37	Cancer, Organ Development, Organismal Injury and Abnormalities

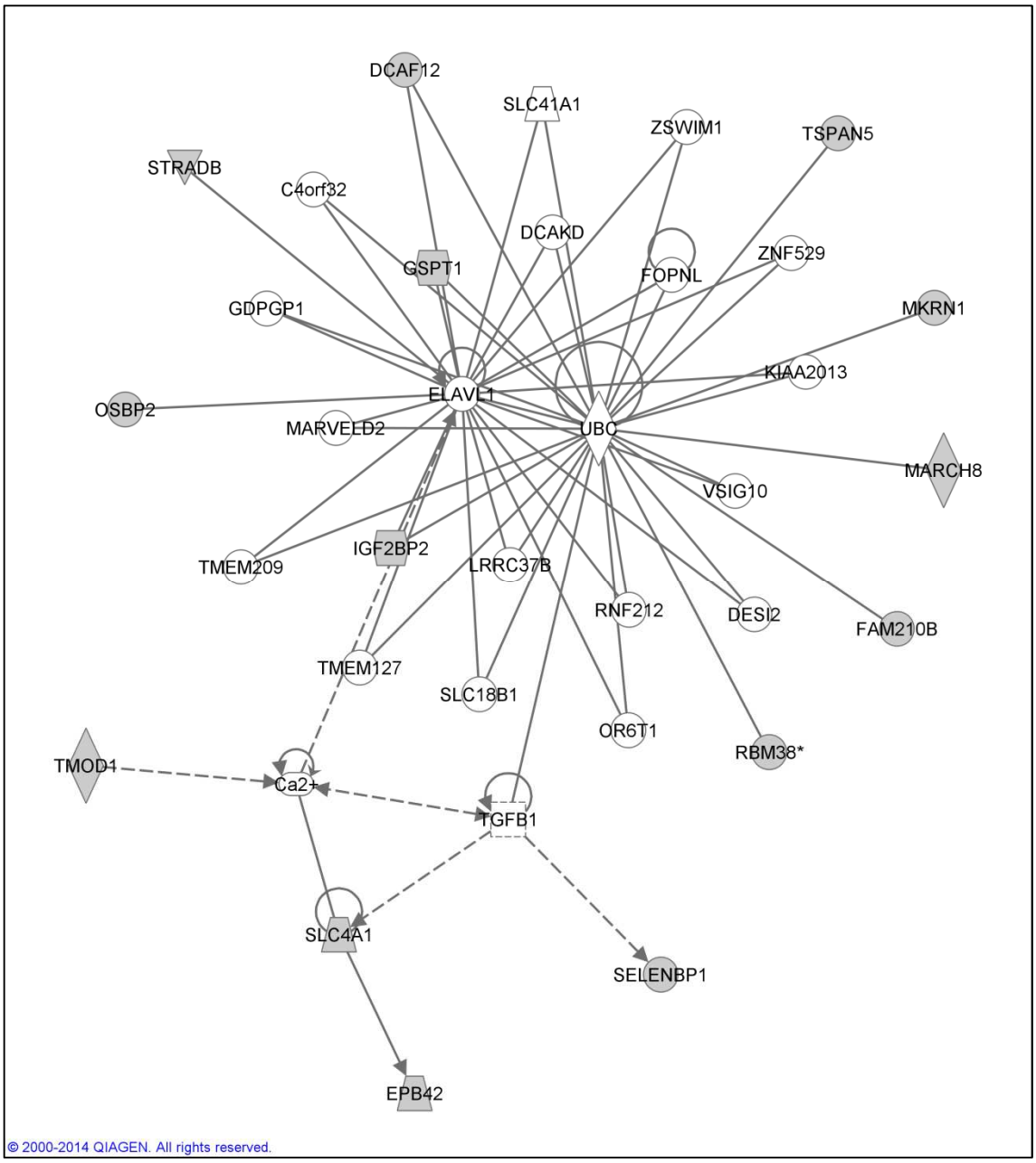

Figure 6-10: Connections between molecules within the IPA network identified using the hub genes in the most significant module in the escitalopram-specific WCNGA analysis. Molecules in grey are hub genes in the WCGNA module, molecules in white are identified as interacting molecules in the Ingenuity Knowledge Base

6.3.2.3 Nortriptyline-specific analysis

When looking at the patients taking nortriptyline, 16 modules of coexpressed gene networks were identified. No significant correlations were observed between change in module eigengene and treatment response. The most significantly correlated module (r=0.36, *p*=0.0228, FDR=0.3873) contains 1,375 gene probes, 1,080 of which were annotated. The correlation between module membership and gene significance was highly significant (r=0.38, p= 6.80×10^{-48}). Pathway analysis of "hub genes" in this module (n=437) showed the best performing

network for enrichment within this module was “Cancer, organ development and organismal injury and abnormalities” (IPA score=48). The details of this IPA network as shown Table 6-7 and Figure 6-11, with additional information of the top 5 networks identified in Appendix F. This is the same network identified in both the whole sample and in the escitalopram-specific sample. Of the 50 hub genes identified in the whole sample analysis, 42 were also identified as hub genes in the nortriptyline-specific analysis.

In the nortriptyline subset of the sample, neither of the two genes identified as significant in the gene-by-gene were included in the WCGNA analysis, as they were not amongst the top 10,000 most connected genes.

Table 6-7: Molecules within the IPA network identified using the hub genes in most significantly associated module in the nortriptyline-specific sample

Analysis	Molecules in Network			Score	Top diseases and functions
	Network eligible		Other		
Nortriptyline	20	DMTN EPB42 GYPC HBD HIVEP3 ICAM4 KLF1 LGALS8 mir-182 MPP1 NCAM1 NPRL3 PCBP2 SH3GLB2 SHARPIN SLC4A1 SSBP3 TRIM10 XPNPEP2 ZSWIM4	ATP1A3 ATXN10 BSG Chymotrypsin CLRN1 Collagen Alpha1 ERK1/2 EXOC1 FBN1 FGF23 Fgfr Haemoglobin PLA2 Spectrin trypsin	48	Cancer, Organ Development, Organismal Injury and Abnormalities

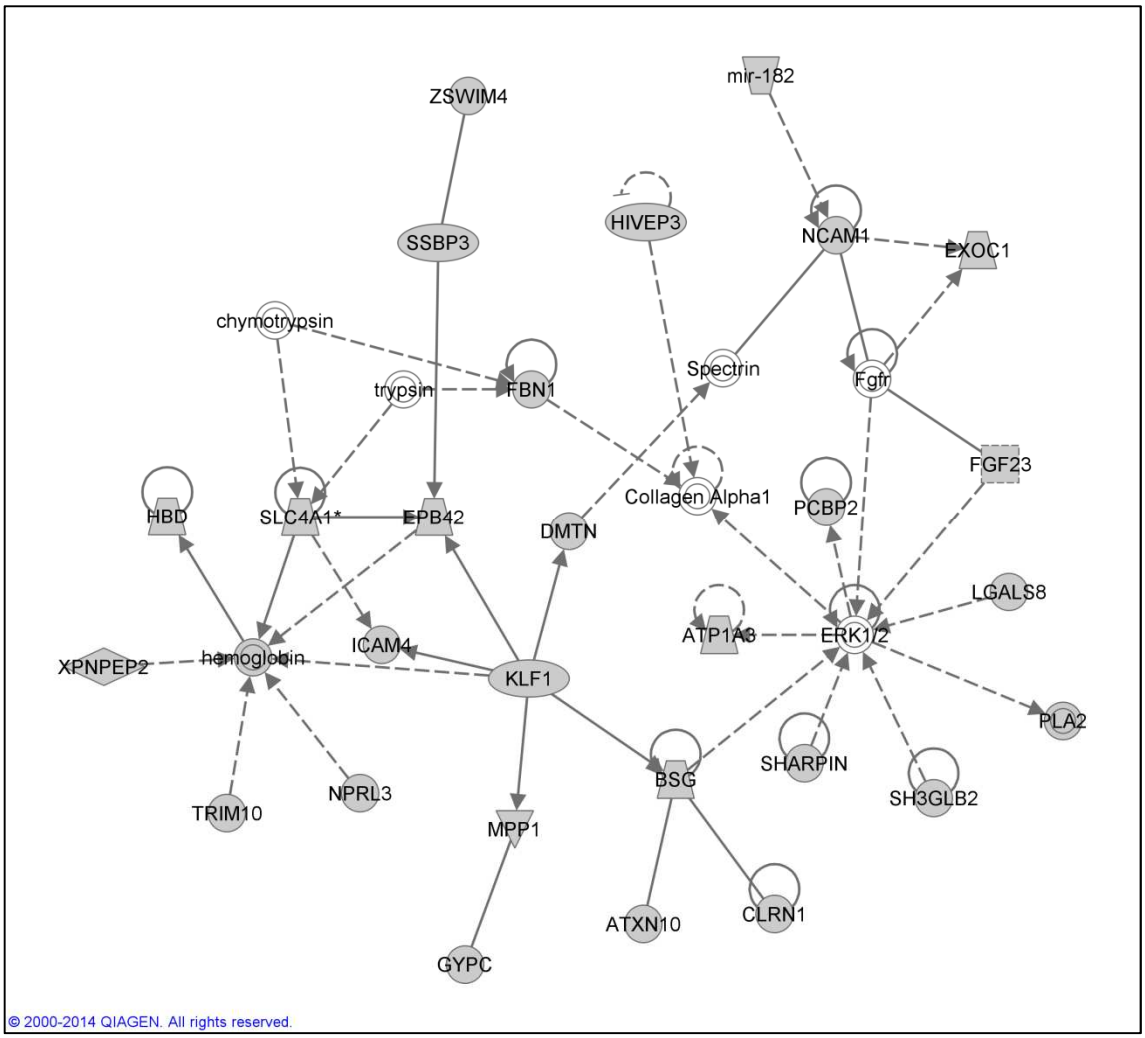

Figure 6-11: Connections between molecules within the IPA network identified using the hub genes in the most significant module in the nortriptyline-specific WCNGA analysis. Molecules in grey are hub genes in the WCGNA module, molecules in white are identified as interacting molecules in the Ingenuity Knowledge Base

6.4 Discussion

6.4.1 Summary of results

6.4.1.1 Individual gene analysis

When employing a gene-by-gene approach, no significant correlations with response were observed for individual genes in either the whole sample or the escitalopram-specific sample. However, significant correlations were observed between changes in gene expression levels and treatment response for two genes, amongst patients taking nortriptyline. This was the smallest subsample to be analysed and should be considered exploratory. The two genes identified were *MMP28* (matrix metalloproteinase 28, involved in extracellular matrix degradation) and *KXD1* (KxDL motif-containing protein 1, involved in endosomal cargo sorting). Furthermore, pathway annotation of those genes reaching suggestive significance in the nortriptyline-specific analysis implicates the IPA network "Cellular Assembly and Organisation, Nervous System Development and Function, Cell Signalling". These annotation terms appear to have direct relevance to previous work on neurotrophic theories of antidepressant action, where efficacy has been linked to second messenger signalling causing gene expression changes and so neurogenesis (Duman *et al*, 1997; Manji *et al*, 2003).

6.4.1.2 Analysis of networks of coexpressed genes

When analysing modules of coexpressed genes, we observed significant correlation between changes in gene expression and treatment response within one module of coexpressed genes in the whole sample analysis. Interestingly, whilst no modules reached significance in the drug-specific analyses, the modules showing strongest association with treatment response in both sets of analyses annotated to the same IPA network terms as that identified in the whole sample analysis (that is "Cancer, organ development and organismal injury and abnormalities"). This shared annotation may indicate that on a systems level, the gene expression changes associated with treatment response are not drug-specific, but instead, affect common biological networks regardless of antidepressant mechanism of action. The link between the networks implicated in this analysis and what is known about antidepressant action is unclear, but it may

be of note that the IPA annotation term “organ development” includes neurogenesis amongst other terms (see Supplementary Materials).

6.4.2 Comparison of methodologies

Whilst the WGCNA-based coexpression approach gives evidence supporting a pattern of common biological pathways between the two drugs considered here, this is not seen when using the individual gene approach. This could indicate that response to different antidepressant medications is not the result of shared action on individual genes, but instead a common action upon genes within particular coexpression networks. However, in the context of the inherent noisiness of microarray data and a high multiple hypothesis testing burden, it may be that this study did not have sufficient statistical power to detect effects using a gene-by-gene approach.

6.4.3 Statistical power

Using G*Power (Faul *et al*, 2007), it is estimated that in the whole sample gene-by-gene analysis, using a threshold of $p<5x10^{-6}$ (reflecting the 29,000 gene probes considered), the sample is large enough to detect correlations of r=0.417 or greater, with 80% power. Therefore, whilst this analysis represents the largest analysis of transcriptomic data from patients taking antidepressants to date, power calculations indicate that there is only sufficient statistical power to detect comparatively large effects using these data. The findings presented here appear to rule out the presence of such large correlations with treatment response for single genes, with the exception of some novel and potentially interesting candidates for the nortriptyline-specific group.

Statistical power is increased using the network approaches of WGCNA. With a threshold of p<0.005 (reflecting the 10 identified coexpression modules), the sample is large enough to detect correlations of r=0.283 or greater, with 80% power.

6.4.4 Comparison to previous literature

The individual gene analysis presented here did not replicate the previous reported association with *IRF7* from a transcriptomic analysis of treatment response of 63 patients with MDD (Mamdani *et al*, 2011). Furthermore those genes which have previously been identified within a smaller subset of GENDEP patients using a candidate gene approach (Cattaneo *et al*, 2013; Powell *et al*, 2013) did not achieve significance when considered within a gene-by-gene, transcriptomic framework. This may reflect the limitations in statistical power to detect smaller effect sizes.

However, when using the more powerful network coexpression approach, the inclusion of neurogenesis in the annotation terms for the modules which are associated with treatment response does align with previous literature highlighting neurogenesis as an important mechanism of antidepressant action (Lee and Kim, 2010). Furthermore, candidate studies (including those in GENDEP) have linked expression changes in genes linked to neuroplasticity to treatment response in patients (Belzeaux *et al*, 2010; Cattaneo *et al*, 2013; Iga *et al*, 2007b). Still, there are other annotation terms within this category which do not appear to link to what is known about the action of antidepressants.

Therefore, further replication is needed to establish the robustness of our findings, as well as extend this work to consider other treatments for depression beyond escitalopram and nortriptyline.

6.4.5 Potential role of placebo effects

In addition to addressing whether these results generalise across different treatments, the role of placebo effects should also be considered. The absence of a placebo arm within GENDEP was necessary in order to make the project more inclusive and open to a wider proportion of patients. Nevertheless, it means that in this analysis it is not possible to disentangle whether those gene expression changes that correlate with treatment response represent a signature of antidepressant action, or symptom improvement.

6.4.6 Tissue specificity

The degree to which data from this study can be used to understand the relationship between gene expression and antidepressant response must, of course, be considered in the context of tissue specificity. Here, blood is being used as a proxy for the key tissue of interest in antidepressant research; the brain. Studies exploring the degree of gene coexpression in blood and brain in humans suggest moderate correlation (Cai *et al*, 2010; Liew *et al*, 2006; Lunnon *et al*, 2012; Sullivan *et al*, 2006), and so whilst blood does appear to be an accessible and useful tissue by which we can probe the transcriptome, the issue of tissue-specificity of gene expression (and indeed drug-induced changes to these gene expression patterns) does constrain the interpretability of our findings.

Nevertheless, the use of blood samples in this study is critical, enabling samples to be taken from patients both before and after treatment for analysis of changes in gene expression patterns.

6.4.7 Conclusions

Overall, we have shown that a relationship between changes in gene expression and treatment response can be observed in patients taking antidepressant medication. The pattern of results indicates these changes take the form of a number of smaller changes acting across a network of coexpressed genes, rather than single genes showing large changes in gene expression levels. Furthermore, even when considering two antidepressants with divergent mechanisms of action, the implicated coexpression networks share the same biological function indicating the changes in gene expression that are associated with treatment response are not drug-specific.

Chapter 7 Examining the genetic control of gene expression to understand predictors of antidepressant response

7.1 Introduction

7.1.1 Understanding the mechanisms of genetic effects

The work that I have presented in this thesis thus far focusses on identifying genetic biomarkers associated with antidepressant treatment. In this chapter, I will move to investigating the downstream consequences of genetic variation, in order to consider the molecular pathways through which genetics might ultimately impact on treatment outcomes.

As noted in the introduction, there are a growing number of identified trait-associated genetic variants (Hindorff *et al*), but we are frequently unable to trace how these identified variants exert their effects. This is of particular interest given that the majority of trait-associated variants identified to date are neither non-synomymous nor missense mutations and so do not directly affect protein structure.

One approach to understanding the mechanisms by which genetic variants act is to consider the impact of genotype on levels of gene expression, within an expression quantitative trait loci (eQTL) framework (see Introduction for further details).

7.1.2 Using eQTLs to interpret genetic and transcriptomic associations

eQTLs are widespread throughout the genome, and have already been shown to be a valuable starting point to unpick the path between genotype and disease, by demonstrating how genetic variation can alter the transcriptome. For example, eQTL analysis of the intronic variants within the *FTO* gene that have been linked to obesity phenotypes demonstrates that the SNPs affect *IRX3* (but not *FTO*) transcript levels; follow-up experiments using knock-out models in mice further found *IRX3* to be associated with weight regulation (Smemo *et al*, 2014).

7.1.3 Using eQTLs to identify genetic associations

eQTLs are not only of value in understanding the mechanism by which previously identified genetic variants exert their effects, they can also be of value when trying to identify those variants.

One of the primary challenges in conducting genome-wide association studies is the collection of sample sizes with sufficient statistical power to detect trait-associated variants, in the face of small genetic effect sizes and a high multiple hypothesis testing burden. Limitations in power, combined with a polygenic signal mean that within each genome-wide association study, there are likely to be a number of SNPs which are truly associated with the trait of interest, but yet fail to reach the threshold for statistical significance.

This issue is particularly pertinent when considering the genetic predictors of antidepressant treatment outcomes. As previously noted, treatment response is under genetic influence (Tansey *et al*, 2013) but the largest mega-analysis to date (n=2,256) failed to identify any variants reaching genome-wide significance (GENDEP Investigators; MARS Investigators; STAR*D Investigators, 2013). The collection of larger sample sizes is needed to gain sufficient power to reach the threshold for genome-wide significance, but pharmacogenetic cohorts are very expensive to collect, given the need for clinical monitoring across the course of treatment.

However, it may be possible to use eQTLs as an annotation tool to aid the extraction of information from those SNPs reaching only suggestive significance levels. There is an enrichment of eQTLs amongst identified trait-associated variants (Nicolae *et al*, 2010), and it has been demonstrated that this fact can be usefully exploited within a Bayesian framework by applying weightings to identified *cis*-eQTL SNPs, to increase discovery of trait-associated variants when performing a GWAS (Knight *et al*, 2011; Li *et al*, 2013). Therefore, using eQTL annotation with the SNPs showing suggestive association with treatment response may be a useful method by which to prioritize variants that are most likely to be causally-linked with the trait.

7.1.4 Using eQTLs to annotate transcriptomic findings

EQTLs can also be used as an annotation tool to aid the interpretation of transcriptomic analyses, by identifying genotypes that may drive transcriptomic patterns. This has been demonstrated in the case of systemic lupus erythematosus, where the disease-linked transcriptomic profile of altered expression levels for *C1QB* and five interferon genes was shown to be controlled by a *trans*-acting eQTL SNP, which had also been previously linked to the disease (Westra *et al*, 2013).

Analyses in GENDEP looking at the transcriptome prior to treatment has highlighted a number of suggestive predictors of treatment response (Tansey *et al*, in prep) and eQTLs can be used to determine if these transcriptomic markers are under genetic control.

7.1.5 Context specificity of eQTLs

Finally, the influence that the phenotype of antidepressant treatment response may have on the genetic control of gene expression can also be explored. As our understanding of the genetic regulation of gene expression develops, it becomes increasingly clear that eQTLs are variable dependent upon context; for example, as discussed in the Introduction, there is a growing body of research aimed at characterising how eQTLs vary between tissues (Grundberg *et al*, 2012; GTEx Consortium, 2013). There is also evidence that the genetic control of gene expression may also be phenotype-dependent. This has most robustly been indicated using the cellular phenotype of immune activation (Fairfax *et al*, 2014), but age- and sex-specific eQTLs have also been identified (Glass *et al*, 2013; Yao *et al*, 2014).

In the same way that the presence of an eQTL may be dependent on levels of immune activation or age, they may also be dependent on treatment response. If this is shown to be the case, and gene expression levels are determined by an interaction between genotype and treatment response, this would indicate that the mechanism by which a genetic variant influences expression levels varies in a manner which is associated with treatment response.

This provides a window into the biological differences that might underlie the variability we observe in treatment outcomes.

7.1.6 Aims of this study

This study aims to explore the pattern of eQTLs observed in the GENDEP dataset, and how these can be used to understand the phenotype of antidepressant response.

Firstly, the study will characterise the eQTLs observed in the GENDEP sample. Overlap with previously published blood eQTL datasets will be assessed, and the enrichment of trait-associated genetic variants amongst eQTLs will be explored.

Secondly, eQTLs identified here will be used to annotate those genetic variants reaching suggestive significance in a previously published GWAS of antidepressant treatment response in the GENDEP sample (Uher *et al*, 2010), in order to highlight variants which may be considered more likely to be causally associated with treatment response.

Thirdly, the presence of eQTLs controlling the expression levels of transcripts that have been identified as possible predictors of treatment response (Tansey et al, in prep) will be explored, to identify whether the response-predictive differences in expression levels might be linked to genotypic variation.

Finally, the presence of conditional eQTLs will be considered, whereby the genetic effect on expression levels is dependent on the phenotype of treatment response.

7.2 Methods

7.2.1 Participants

Participants included in this chapter are drawn from the GENDEP sample, as described in detail in Chapter 2.2. Figure 7-1 shows the patients included in analyses presented in this chapter.

7.2.2 Treatment response

Response to antidepressant treatment was measured using percentage change in MADRS score from week zero to week twelve of the study. This measure was adjusted for age and recruitment centre (details in Chapter 2.4.1).

7.2.3 Genomic data

Genome-wide genotypic data was obtained for 796 patients within the GENDEP sample using the Illumina Human610-quad bead chip (Illumina Inc., San Diego). Standard quality control measures were applied, population stratification principal components calculated and imputation using 1000 Genomes data was undertaken, as described in Chapter 2.6.2.

7.2.4 Transcriptomic data

Gene expression data was collected at week zero (prior to antidepressant treatment) from blood samples, using Illumina HumanHT-12 v4 Expression BeadChip microarrays (Illumina Inc., San Diego). Standard quality control measures and batch correction were applied, as described in Chapter 2.7

Aligning the genomic and transcriptomic data, information on 8,317,505 SNPs and 27,187 gene probes was available for 207 patients.

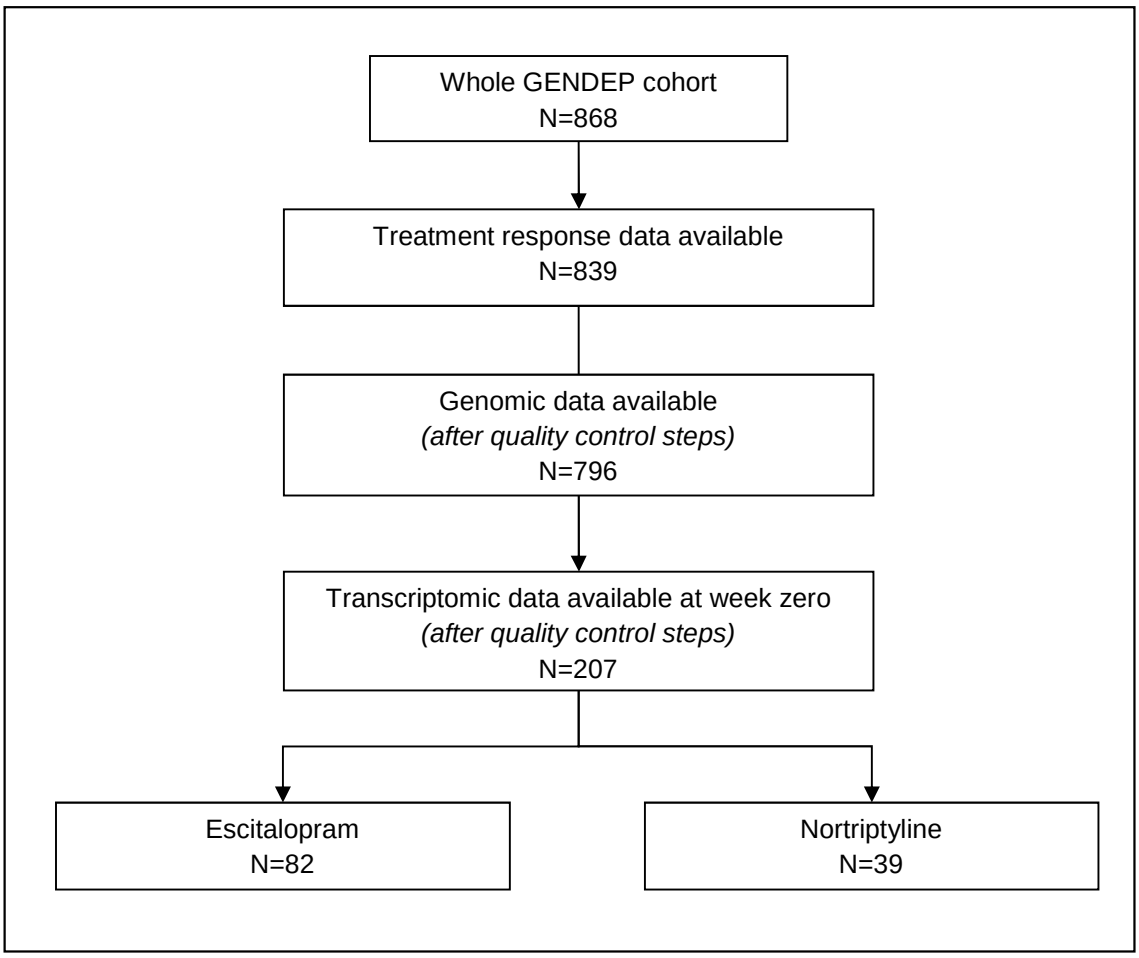

Figure 7-1: Sample included in Chapter 7

7.2.5 Statistical Analysis

7.2.5.1 Identification of eQTLs

eQTL analysis was performed using the R package Matrix eQTL (Shabalin, 2012). Prior to analysis, gene expression values were transformed to correct for outliers, using the protocol employed by the Genotype-Tissue expression (GTex) project (http://www.gtexportal.org/). For each gene, expression values were ranked across samples then mapped onto a standard normal distribution.

Covariates entered into the model included the first four principal components calculated from the genotypic data (to account for population stratification), the first principal component calculated from the gene expression data, as surrogate measures of unwanted variation in gene expression data from non-genetic causes (Biswas *et al*, 2008; Leek and Storey, 2007),

estimated proportions of cell type (as calculated using CellMix (Gaujoux *et al*, 2013), see Chapter 2.7 for details), age, sex and treatment response.

An eQTL analysis involves performing a genome-wide association study for each of the 27,187 probes. To correct for the number of tests performed, the threshold of genome-wide significance ($p<5x10^{-8}$) was divided by the number of probes tested, to give a threshold of $p<1.84x10^{-13}$. This represents a conservative Bonferroni adjustment, as the correlated structure of the transcriptomic data has not been considered. However, the stringent threshold was used, given the issue of limited replication within the eQTL literature (Breitling *et al*, 2008). Suggestive findings were defined as those where $p<1x10^{-8}$. Results were then divided into *cis* and *trans* eQTLs; *cis*-eQTLs were defined as those within 500KB of the gene probe, all other associations were designated as *trans*-eQTLs. Any gene probes containing polymorphisms of a minor allele frequency greater than 0.01 (within the 1000 genomes sample of European ancestry) were removed from the analysis using BEDtools (Quinlan and Hall, 2010) and the protocol described by Ramasamy *et al* (2013). Probe locations were determined using the illuminaHumanv4.db Bioconductor package in R (Dunning *et al*), and any probes mapping to multiple locations were removed.

To obtain the number of independent associations, for each gene probe with more than one SNP significantly associated, linkage disequilibrium clumping was performed using PLINKv1.07 (http://pngu.mgh.harvard.edu/purcell/plink/) (Purcell *et al*, 2007).

To consider how reliable the eQTLs in this analysis are, they were compared with those previously identified in peripheral blood samples taken from 5,311 individuals (Westra *et al*, 2013). The authors of this large sample (referred to hereafter as the Westra dataset) made all *cis* and *trans*-eQTLs that reached an FDR<0.5 publicly available (http://genenetwork.nl/bloodeqtlbrowser/). It is of note that the Westra dataset used gene probes included on the Illumina Human HT12v3 array (these overlap but do not completely match those included on the v4 array used in this study), and that for the *trans*-eQTL analysis, only SNPs

that were included in the NHGRI catalogue of published genome-wide association studies (http://www.genome.gov/gwastudies/) were entered into the analysis. An eQTL was only considered to be replicated in the Westra dataset if the exact combination of SNP and probe was observed, with the same direction of effect.

7.2.5.2 Assessing eQTL enrichment amongst genetic associations for any trait

Using the NHGRI catalogue of published genome-wide association studies (http://www.genome.gov/gwastudies/; accessed 06/05/14), a list of all previously published genetic associations passing a threshold of $p<5x10^{-8}$ was obtained, containing a total of 4,695 SNPs. Using PLINK, additional tagging SNPs (with an r2 of 0.5 or greater and within 250kb of the index SNP) were also identified to create a list of trait-associated genetic variants (143,375 SNPs). We then assessed overlap between these trait-associated SNPs and the *cis*-eQTL loci.

To assess whether the number of trait-associated SNPs which were also *cis*-eQTLs is greater than would be expected by chance, permutation tests were run. A random sample of SNPs was selected, equal in size to the number of *cis*-eQTL loci observed. The overlap between this random sample of SNPs and the trait-associated variants (all GWAS hits plus tagging SNPs) was assessed. This was repeated 1,000 times to obtain a null distribution of the expected number of *cis*-eQTLs that would be observed by chance, and used to assess enrichment.

7.2.5.3 Assessing eQTL enrichment amongst genetic associations with antidepressant response

To then explore eQTLs amongst suggestive genetic associations with treatment response, a list of SNPs reaching suggestive levels in significance (60 SNPs $p<0.0001$) in the previously published GWAS in GENDEP (Uher *et al*, 2010) was obtained. Additional tagging SNPs were identified using PLINK (as described above), creating a total list of 1154 response-associated SNPs, to compare with the identified *cis*-eQTL loci.

7.2.5.4 Using eQTLs to explore transcriptomic predictors of treatment response

Using the results from the transcriptomic analysis of baseline predictors of treatment response in GENDEP (Tansey et al, in prep), we explored whether any of the genes showing suggestive association with response ($p<0.001$, $n=53$) were also identified here as eQTL genes.

7.2.5.5 Identifying eQTLs conditional on treatment response

Finally, a second eQTL analysis was performed to explore whether it was possible to identify eQTLs that are conditional on treatment response outcomes. In this analysis, rather than covarying for treatment response, the interaction between treatment response and genotype was tested for association with gene expression levels.

7.3 Results

7.3.1.1 Identification of eQTLs

After removing all eQTLs involving probes containing polymorphisms, a total of 25,269 significant *cis*-eQTLs were observed. This represents 668 independent signals, affecting 444 gene probes. The top 20 independent signals are shown in Table 7-1. The mean effect size of *cis* eQTLs was R^2=0.375, with a maximum effect size of R^2=0.769. Moving to a suggestive level of significance ($p<1x10^{-8}$), an additional 32,845 *cis*-eQTLs were detected. A Manhattan plot showing all *cis*-eQTL signals is shown in Figure 7-2.

Focussing on the *trans*-eQTLs, 5,898 significant SNP-gene probe associations were observed, representing 103 independent signals affecting 49 gene probes. The maximum number of gene probes associated with any single SNP was 12. The mean affect size of the *trans*-eQTLs was R^2=0.403 (maximum effect size of R^2=0.760). At a suggestive level of significance, an additional 3,504 *trans*-eQTLs were observed.

In line with previous studies (Dimas et al, 2009; Stranger et al, 2007; Veyrieras et al, 2008), we also noted that the effect sizes for *cis*-eQTLs are larger the closer they are to the gene probe (Figure 7-3).

7.3.2 Replication of eQTL signals

Of the 25,269 *cis*-eQTLs observed, 4,827 were also detected in the Westra dataset. No replications of the *trans*-eQTLs were detected, although it should be noted that whilst in the current study, all available variants were tested, Westra *et al* (2013) limited their *trans*-eQTL analysis to 4,542 variants included in the NHGRI GWAS catalogue.

Table 7-1: The 20 most signficant *cis*-eQTLs (SNP annotation from snp-nexus.org)

SNP	Gene Name	Illumina Probe	Chr	P value	Beta	R^2	P value in Westra dataset	SNP Annotation
rs1131017	*RPS26*	ILMN_3299955	12	1.48E-63	1.26	0.77	-	5UTR
rs1131017	*RPS26*	ILMN_1695585	12	8.80E-60	1.22	0.75	-	5UTR
rs3813976	-	ILMN_1818577	1	2.90E-59	1.13	0.74	-	5upstream
rs10760117	*PSMD5-AS1*	ILMN_3236498	9	2.66E-58	1.23	0.74	-	intronic
rs1131017	*RPS26*	ILMN_3209193	12	1.01E-53	1.19	0.71	-	5UTR
rs11717719	*CHST13*	ILMN_1734707	3	3.52E-51	-1.13	0.69	-	intronic
rs1131017	*RPS26P31*	ILMN_3285153	12	2.22E-49	1.18	0.68	-	-
rs10239340	*IRF5*	ILMN_2312606	7	5.67E-49	1.13	0.67	9.81E-198	5upstream
rs9471975	*PEX6*	ILMN_1683279	6	9.77E-49	-1.18	0.67	-	-
rs9890200	*SPATA20*	ILMN_1687247	17	4.70E-47	-1.14	0.66	-	5UTR
rs1131017	*LOC641768*	ILMN_3242288	12	5.86E-47	1.17	0.66	-	-
rs2431529	*PAM*	ILMN_2313901	5	1.00E-46	1.18	0.66	9.81E-198	-
rs61913527	*CLEC12A*	ILMN_1663142	12	3.82E-45	1.20	0.64	-	intronic
rs12933746	*LPCAT2*	ILMN_1796335	16	1.06E-44	-1.05	0.64	-	3UTR
rs7309256	*CLEC12A*	ILMN_1711453	12	1.98E-44	1.24	0.64	9.81E-198	-
rs4731533	*IRF5*	ILMN_1670576	7	2.63E-44	1.10	0.63	-	5upstream
rs4766601	*KCTD10*	ILMN_1719064	12	1.88E-43	1.32	0.63	9.81E-198	intronic
rs11191667	*USMG5*	ILMN_1773313	10	2.20E-43	1.14	0.63	-	-
rs1045599	*ZSWIM7*	ILMN_3298167	17	3.87E-43	-1.12	0.62	-	3UTR
rs2709398	*METTL21A*	ILMN_3250243	2	1.04E-42	-1.34	0.62	-	-

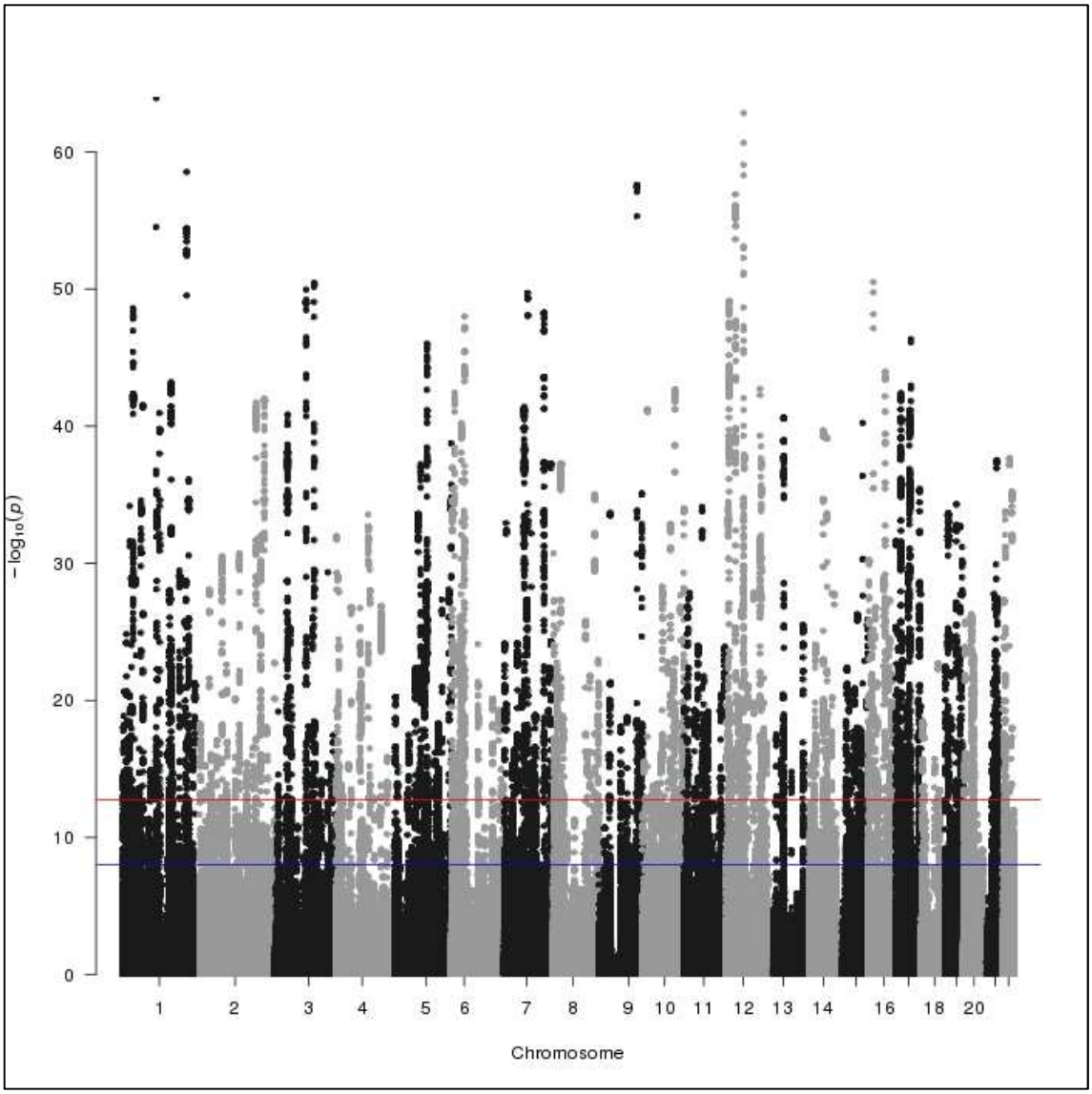

Figure 7-2: Manhattan plot of all *cis*-eQTLs.

Red line indicates threshold for significance (p<1.84x10-13), blue line indicates threshold for suggestive findings (p<1x10-8)

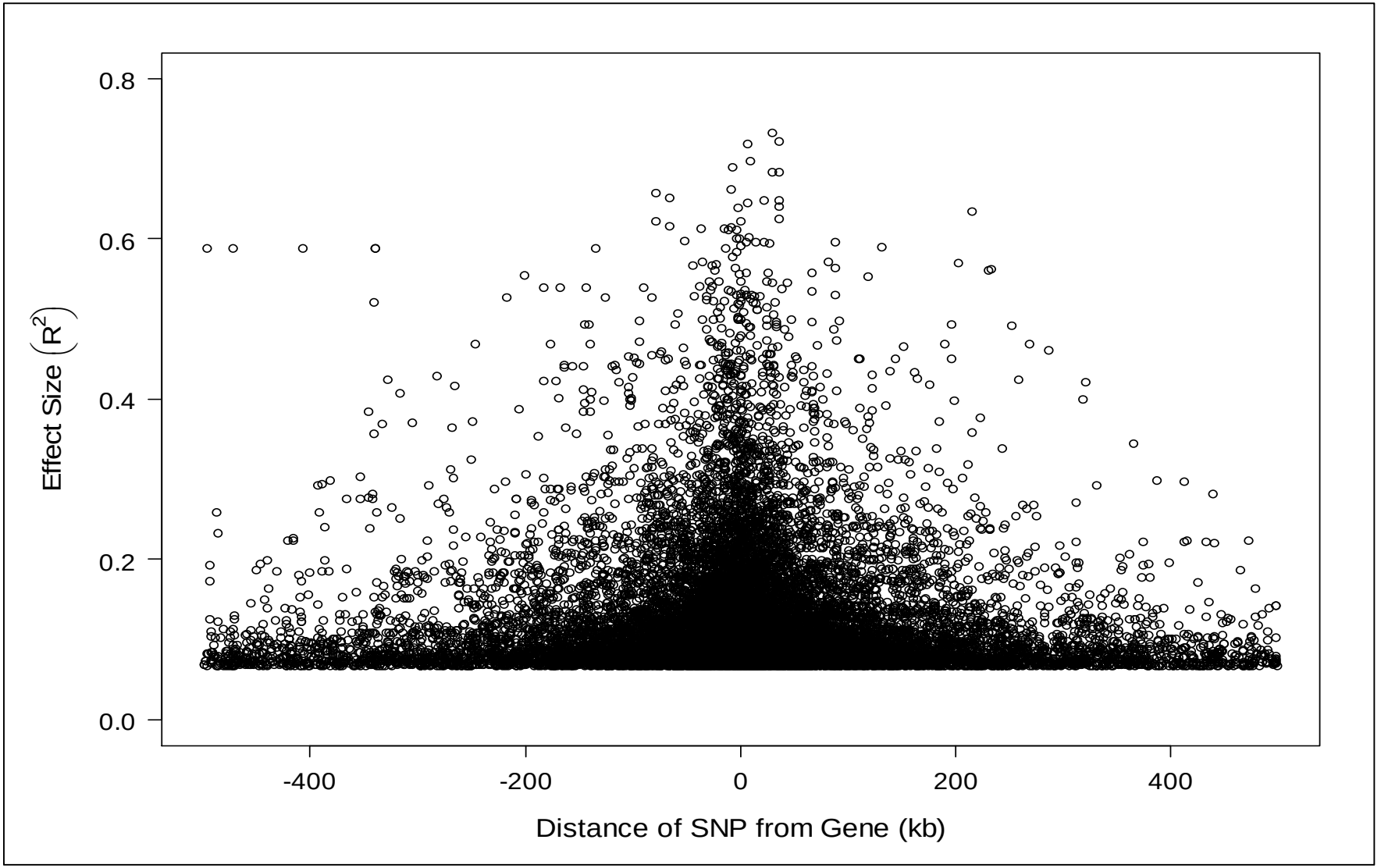

Figure 7-3: The distance between the cis-eQTL SNP and gene, in relation to the effect size of the eQTL

7.3.2.1 Enrichment of eQTL-SNPs amongst GWAS hits associated with any trait

To assess whether the cis-eQTLs identified in this analysis were enriched amongst GWAS-significant variants, a total of 143,375 trait associated SNPs (4,695 from the NHGRI GWAS catalogue plus tagging SNPs) were compared with the 668 *cis*-eQTL loci identified in this analysis. A total of 77 SNPs were present in both lists and permutation tests show this demonstrates significant enrichment (p<0.001, see Figure 7-4)

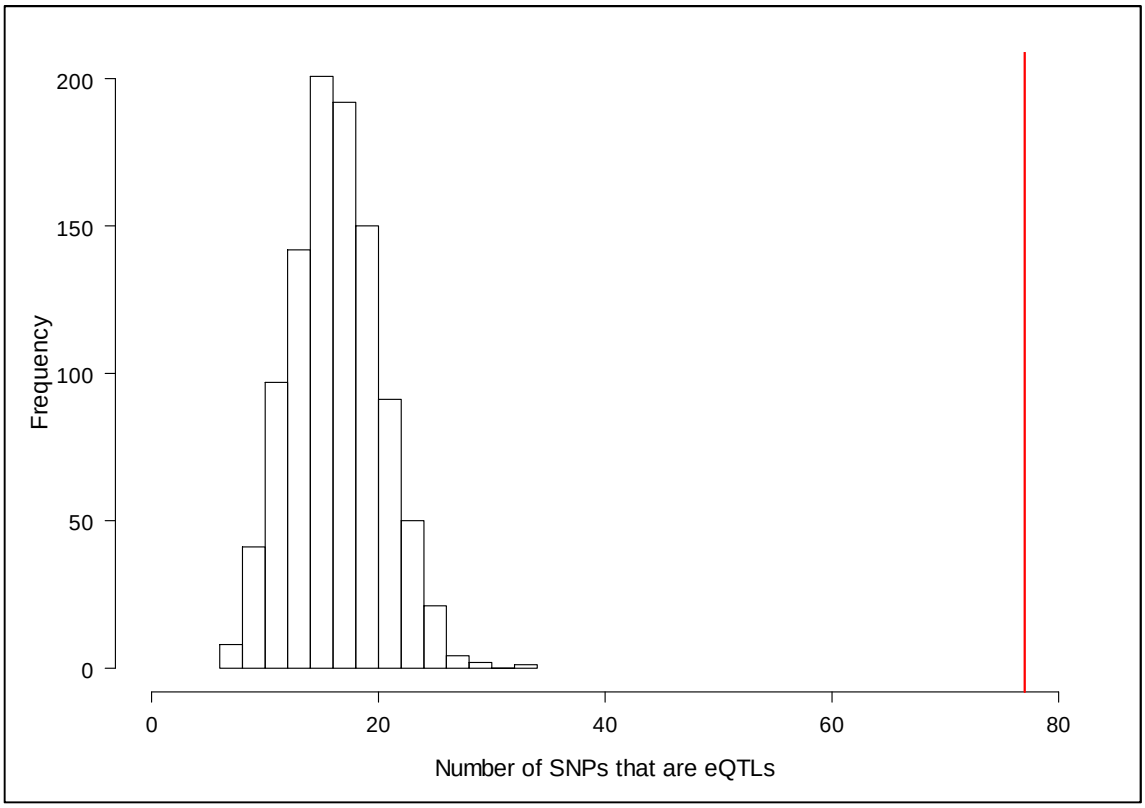

Figure 7-4: Enrichment of *cis*-eQTLs amongst any trait-associated SNPs. Null distribution from 1,000 permutations is shown, with a red line indicating the number of cis-eQTL loci observed amongst trait-associated SNPs

7.3.2.2 eQTL-SNPs amongst suggestive findings linked to antidepressant response

Considering the suggestive response-associated SNPs (n=1145; 60 identified SNPs plus tagging variants) identified in the GWAS of treatment response in GENDEP (Uher *et al*, 2010), 1 SNP (rs10747570) was located within a *cis*-eQTL locus. This locus was located on chromosome 12, and associated with expression levels of the gene *CERS5*, which encodes Ceramide Synthase 5. The *cis*-eQTL locus contains 56 SNPs in high LD, 23 of these are replicated in the Westra dataset. Furthermore, this *cis*-eQTL locus is also reported in tibial nerve tissue in the GTex dataset (http://www.gtexportal.org/home/).

The pattern of high LD in this locus (shown in Appendix G) means that it is difficult to identify the specific SNP involved, and so the potential mechanism underlying this eQTL. Nevertheless, Figure 7-5 shows the context of this eQTL locus, with the position of all implicated SNPs, the CERS5 gene and details from the ENCODE project on the regulation of transcription (see details of ENCODE Integrated Regulation Tracks on the UCSC Browser for further details, www.genome.ucsc.edu). Table 7-2 to Table 7-4 show annotations that have been identified for the 56 identified SNPs in the eQTL locus, using the HaploReg tool (www.broadinstitute .org/mammals/haploreg/haploreg.php; (Ward and Kellis, 2012)

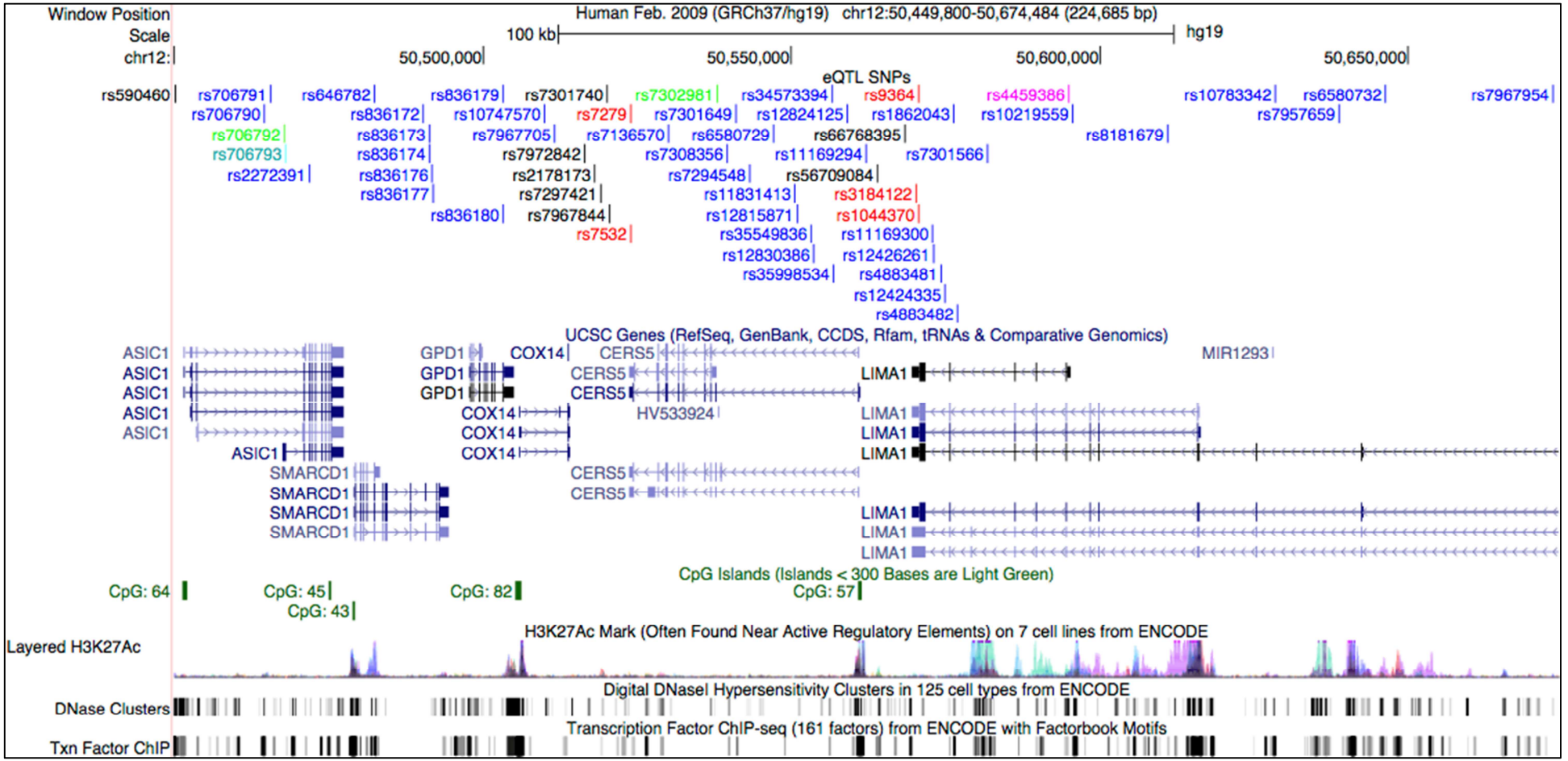

Figure 7-5: Position of SNPs within eQTL locus affecting CERS5 expression. . (SNPs coloured by known function; blue = intronic, red = 3'UTR, magenta,=5'UTR, green = missense, cyan=synonymous, black=intergenic

Table 7-2: Annotations of 56 SNPs identified in CERS5 eQTL locus, as identified using Haploreg v2 (A)

Variant	Allele	EUR freq	Promoter histone marks	Enhancer histone marks	DNAse	Proteins bound	eQTL tissues	Motifs changed	GENCODE gene	dbSNP funct annot.
rs10747570	A/G	0.59							COX14	intronic
rs7967705	T/C	0.59						AhR, EBF, Pax-5	COX14	intronic
rs7972842	G/T	0.6			WERI-Rb-1			CDP, Cdx	COX14	
rs2178173	A/G	0.6							COX14	
rs7297421	C/T	0.6		GM12878				Foxa, Foxp3	COX14	
rs7301740	C/G	0.6						AhR, EBF	COX14	
rs7967844	G/C	0.59						4 altered motifs	COX14	
rs7136570	A/G	0.6						Bcl6b, Pax-4	COX14	intronic
rs7532	T/C	0.6						Ets, Hic1, Maf	COX14	3'-UTR
rs7279	A/G	0.6						HDAC2, Irf, p300	COX14	3'-UTR
rs836179	A/G	0.4		GM12878, K562, HepG2	HUVEC, HeLa-S3, GM19238			4 altered motifs	GPD1	intronic
rs836180	C/T	0.41		GM12878, K562, HepG2	A549			5 altered motifs	GPD1	synon.
rs6580729	G/A	0.4						5 altered motifs	COX14	intronic
rs11831413	A/T	0.4						8 altered motifs	COX14	intronic
rs7294548	A/G	0.6		Huvec	4 cell types			AP-4, Ascl2, YY1	COX14	intronic
rs12815871	G/A	0.4							COX14	intronic
rs7301649	T/C	0.6						Hltf	COX14	intronic
rs35549836	C/T	0.4							RP11-411N4.1	intronic
rs12830386	G/A	0.4						CEBPB, CEBPD, GR	RP11-411N4.1	intronic
rs12824125	A/G	0.4		6 cell types				5 altered motifs	COX14	intronic
rs34573394	C/A	0.4						7 altered motifs	RP11-411N4.1	intronic

Table 7-3: Annotations of 56 SNPs identified in CERS5 eQTL locus, as identified using Haploreg v2 (B).

Variant	Allele	EUR freq	Promoter histone marks	Enhancer histone marks	DNAse	Proteins bound	eQTL tissues	Motifs changed	GENCODE gene	dbSNP funct annot.
rs35998534	A/G	0.4						LF-A1, TATA, YY1	COX14	intronic
rs7302981	A/G	0.6							COX14	missense
rs7308356	A/G	0.6						NRSF	COX14	intronic
rs11169300	T/C	0.59						4 altered motifs	LIMA1	intronic
rs12426261	A/G	0.6						4 altered motifs	LIMA1	intronic
rs4883481	T/C	0.6						GCNF, Hand1, Nr2f2	LIMA1	intronic
rs1862043	G/C	0.59			HRGEC, HRPEpiC, SK-N-MC	SMC3			LIMA1	intronic
rs4883482	A/G	0.59						7 altered motifs	LIMA1	intronic
rs11169294	A/G	0.4	HepG2	6 cell types	pHTE			Ets	709bp 5' of CERS5	
rs706790	A/G	0.43	HSMM	K562, Huvec, NHEK	116 cell types	11 bound proteins		5 altered motifs	ASIC1	intronic
rs2272391	A/G	0.54						6 altered motifs	ASIC1	intronic
rs706792	G/T	0.44						10 altered motifs	ASIC1	missense
rs836172	C/G	0.44							SMARCD1	intronic
rs836177	A/G	0.44			WI-38			Zfp691	SMARCD1	intronic
rs836174	G/A	0.44						Brachyury, LUN-1, Nkx2	SMARCD1	intronic
rs836173	C/T	0.44						YY1	SMARCD1	intronic
rs836176	C/T	0.44						4 altered motifs	SMARCD1	intronic
rs706793	G/A	0.44						4 altered motifs	ASIC1	synon.
rs646782	T/C	0.44		4 cell types				5 altered motifs	SMARCD1	intronic
rs706791	A/G	0.44		K562				Pax-4, Pax-5, Sin3Ak-20	ASIC1	intronic
rs1044370	T/C	0.44						8 altered motifs	LIMA1	3'-UTR

Table 7-4: Annotations of 56 SNPs identified in CERS5 eQTL locus, as identified using Haploreg v2 (C)

Variant	Allele	EUR freq	Promoter histone marks	Enhancer histone marks	DNAse	Proteins bound	eQTL tissues	Motifs changed	GENCODE gene	dbSNP funct annot.
rs66768395	G/T	0.44			5 cell types			Myf	1.3kb 3' of LIMA1	
rs56709084	T/C	0.44		H1, HepG2, HSMM		SETDB1			2.6kb 5' of CERS5	
rs7301566	T/C	0.55		7 cell types	HA-sp, Th2	CTCF, POL24H8		7 altered motifs	RP3-405J10.3	intronic
rs4459386	A/G	0.56		4 cell types	37 cell types	FOXA1, GATA3		4 altered motifs	LIMA1	5'-UTR
rs10219559	T/C	0.55		4 cell types				5 altered motifs	LIMA1	intronic
rs8181679	T/C	0.56	GM12878	NHEK, HMEC				4 altered motifs	LIMA1	intronic
rs7967954	G/A	0.56			GM19238, HRCEpiC, RPTEC			4 altered motifs	LIMA1	intronic
rs10783342	T/C	0.56		4 cell types	HeLa-S3	9 bound proteins			LIMA1	intronic
rs6580732	T/G	0.56		K562, NHEK, GM12878					LIMA1	intronic
rs7957659	T/C	0.56		NHLF				CEBPG	LIMA1	intronic
rs590460	T/C	0.66	7 cell types	HMEC	105 cell types	10 bound proteins		4 altered motifs	1.4kb 5' of ASIC1	
rs3184122	A/G	0.34							LIMA1	3'-UTR
rs9364	G/A	0.65						Lhx8	LIMA1	3'-UTR
rs12424335	C/T	0.35						4 altered motifs	LIMA1	intronic

7.3.2.3 Annotation of suggestive transcriptomic signatures linked to antidepressant response

Looking at gene expression levels that had been previously identified as suggestively associated with treatment outcomes in this dataset (Tansey *et al*, in prep), one of the 53 identified gene probes was also a *cis*-eQTL gene in this analysis. The implicated gene was *GSTM3*, which encodes Glutathione S-transferase M3, a brain expressed enzyme linked to the uptake and detoxification of a range of toxins and drugs at the blood brain barrier. Expression levels of the gene were associated with one locus containing 18 *cis*-acting SNPs. Figure 7-6 shows how gene expression levels change with genotype for the most significant SNP at this locus. The *cis* eQTL is not replicated in the Westra dataset, however is observed within the GTEx portal, for a range of tissues considered.

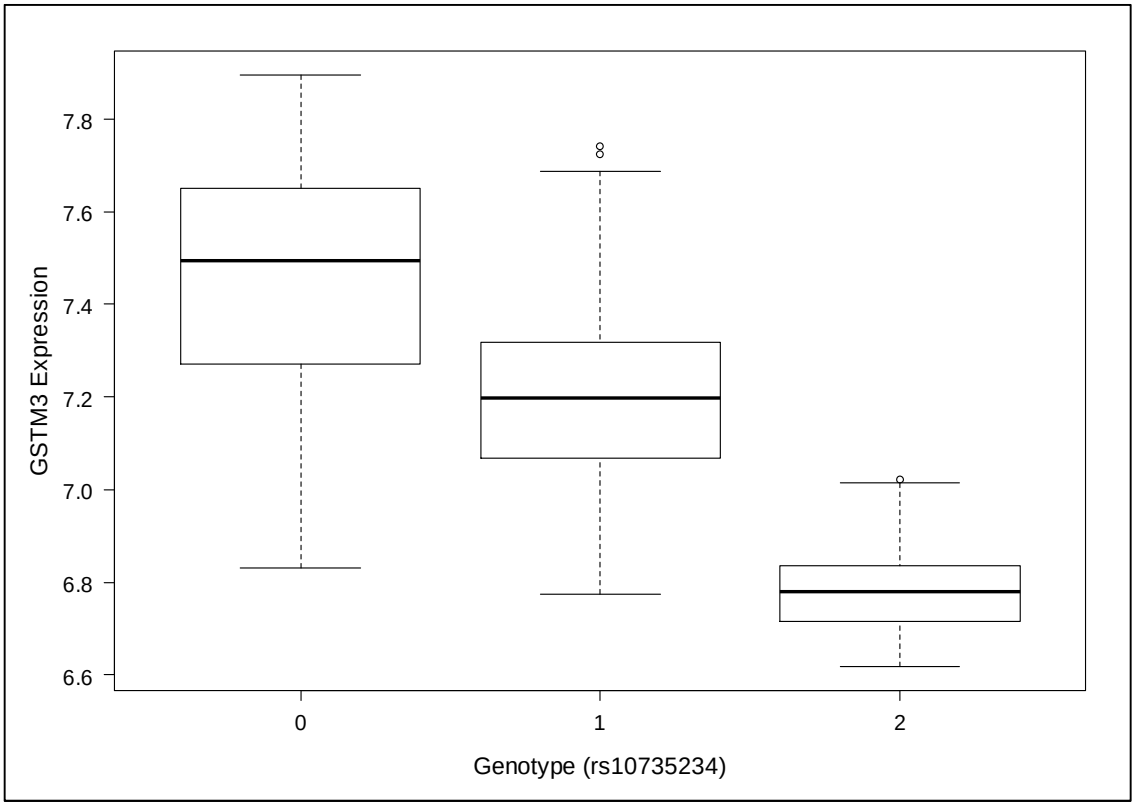

Figure 7-6: Relationship between genotype and expression levels of GSTM3

The high LD within this locus is shown In Appendix G, which makes the identification of the possible mechanism underlying this eQTL difficult to establish. Nevertheless, Figure 7-7 shows the context of the eQTL locus, and Table 7-5 shows the annotations available for the implicated SNPs at this eQTL locus in HaploReg. Interestingly, several of the SNPs have been identified as eQTL SNPs in human brain tissue (Gibbs et al, 2010).

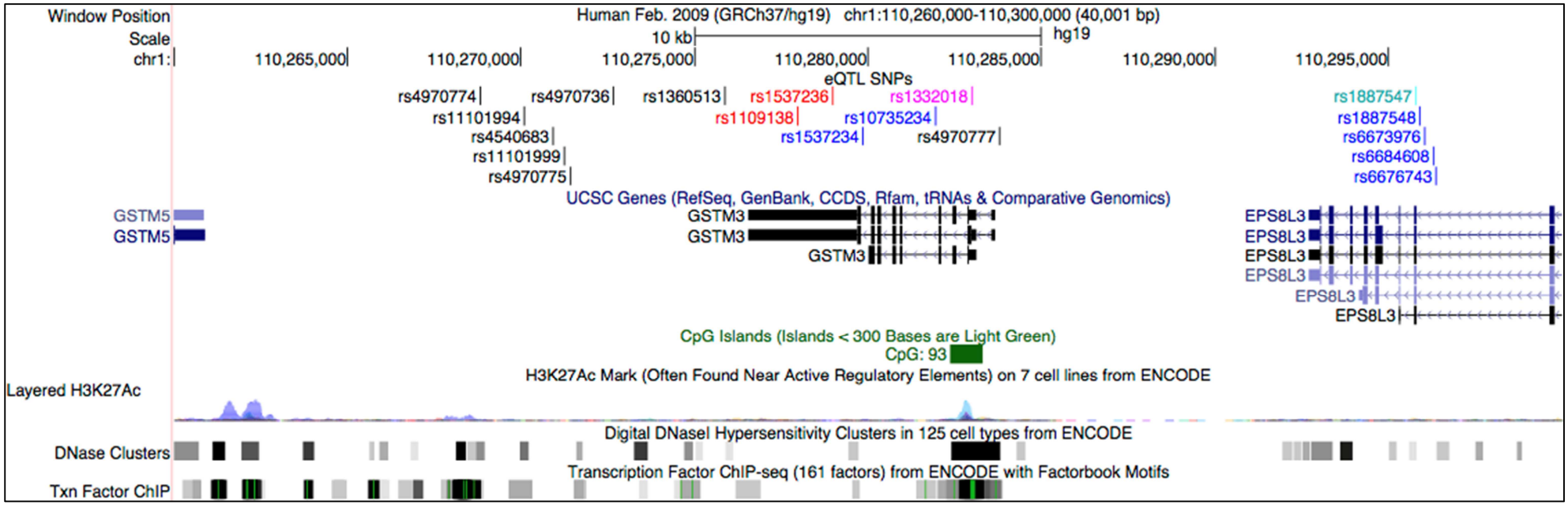

Figure 7-7: Position of SNPs within eQTL locus affecting GSTM3 expression. Top track shows eQTL SNPs identified in eQTL locus controlling expression of GSTM3. SNPs are coloured by known function; blue = intronic, red = 3'UTR, magenta,=5'UTR, green = missense, cyan=synonymous, black=intergenic.

Table 7-5: Annotations of 18 SNPs identified in GSTM3 eQTL locus, as identified using Haploreg v2. eQTL tissues identified; CB= cerebellum, FC= frontal cortex, TC= temporal cortex.

Variant	Allele	EUR freq	Promoter histone marks	Enhancer histone marks	DNAse	Proteins bound	eQTL tissues	Motifs changed	GENCODE gene	dbSNP funct annot.
rs1332018	G/T	0.58	9 cell types		98 cell types	12 bound proteins	CB, FC, Pons, TC	RXRA,SMC3	GSTM3	5'-UTR
rs10735234	G/A	0.59	HSMM	4 cell types	Osteobl		CB, FC, Pons, TC	4 altered motifs	GSTM3	intronic
rs4970777	C/T	0.58		Huvec, HepG2	HPAEC, HRGEC	CTCF	CB, FC, Pons, TC	4 altered motifs	GSTM3	
rs4970736	C/T	0.59						5 altered motifs	RP4-735C1.4	
rs1109138	T/C	0.59					CB, FC, Pons, TC	HEN1	GSTM3	3'-UTR
rs4970775	T/G	0.6						4 altered motifs	RP4-735C1.4	
rs11101994	A/C	0.59			H7-hESC, SK-N-MC			4 altered motifs	RP4-735C1.4	
rs4540683	T/G	0.59						EWSR1-FLI1, HDAC2	RP4-735C1.4	
rs1887548	C/G	0.61			Osteobl		CB	5 altered motifs	EPS8L3	intronic
rs1887547	T/C	0.61			HSMMtube		CB	BATF, Pax-6, RFX5	EPS8L3	synon.
rs6673976	C/T	0.61					CB	Zfp410	EPS8L3	intronic
rs6684608	T/C	0.61					CB	4 altered motifs	EPS8L3	intronic
rs6676743	C/T	0.61					CB	8 altered motifs	EPS8L3	intronic
rs1537236	C/T	0.51					CB, FT, TC	Sox	GSTM3	3'-UTR
rs1360513	T/C	0.47						Maf	RP4-735C1.4	
rs4970774	A/C	0.47		K562	4 cell types	FOXA1	FC, TC	29 altered motifs	RP4-735C1.4	
rs11101999	G/C	0.46						6 altered motifs	RP4-735C1.4	
rs1537234	C/A	0.43						Cdx, Pax-4	GSTM3	intronic

7.3.2.4 Conditional eQTLs

Finally, we explored whether genetic control of gene expression varied with the phenotype of treatment response; that is whether any eQTLs were conditional on the phenotype of treatment response. No conditional eQTLs passed the threshold for significance. However, 801 *trans*-eQTLs reached suggestive levels of significance (representing 154 independent signals affecting 143 probes), with a mean R^2 of 0.165 (maximum R^2=0.206). The top 20 *trans*-eQTL loci are shown in Table 7-6, with the most significant eQTL illustrated in Figure 7-8. It should be noted that for illustration purposes, in the boxplot, the sample has been divided into responders and non-responders, however the analysis used a quantitative phenotype of treatment response.

Using WebGestalt to explore the implicated gene probes in this analysis, no GO categories showed enrichment with a $p<0.05$, except one relating to photoreceptor cell fate commitment ($p=0.0265$, category contains two gene probes).

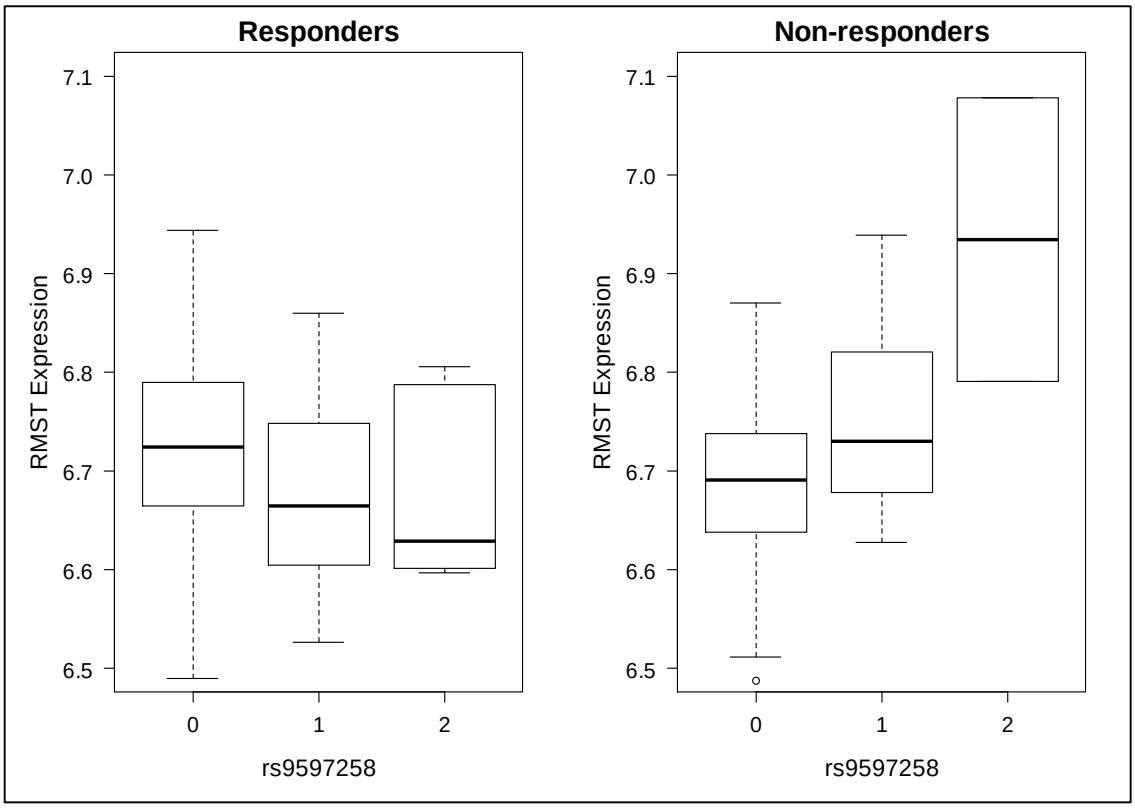

Figure 7-8: Boxplot showing illustrative example of relationship between treatment response, genotype and gene expression levels for the most significant conditional trans-eQTL observed (rs9597258, *RMST* gene probe ILMN_3245464).

Table 7-6: The 20 most significant conditional *trans*-eQTL loci

SNP	SNP C'some	Genetic context of SNP	Gene Name	Illumina Probe	Probe C'some	P value	Beta	R^2
rs9597258	13	Intergenic	*RMST*	ILMN_3245464	12	3.21E-11	-0.03	0.20
rs4742127	9	Intergenic	*ZSWIM4*	ILMN_1735231	19	1.46E-10	-0.02	0.19
rs34693904	4	*COL25A1* (intronic)	*GSDMB*	ILMN_1666206	17	2.28E-10	-0.02	0.19
rs11745163	5	*SPOCK1* (intronic)	*TMPRSS11BNL*	ILMN_3243223	4	3.13E-10	-0.02	0.19
rs78407051	1	Intergenic	*LOC100130133*	ILMN_3231667	Y	4.45E-10	-0.02	0.18
rs2446125	8	*CNBD1* (intronic)	-	ILMN_1865268	16	5.78E-10	-0.02	0.18
rs113338623	7	Intergenic	*SNORA22*	ILMN_3246713	7	6.19E-10	-0.03	0.18
rs1448903	2	*ADAM23* (intronic)	*LOC388114*	ILMN_1669777	15	6.53E-10	-0.03	0.18
rs2021974	20	*MACROD2* (intronic)	*LOC729154*	ILMN_3297582	15	7.02E-10	-0.02	0.18
rs27003	5	*MCTP1* (intronic)	*LOC650154*	ILMN_1773323	5	8.62E-10	-0.02	0.18
rs2418402	9	Intergenic	*LOC402617*	ILMN_1708542	7	9.90E-10	0.03	0.18
rs7814809	8	Intergenic	-	ILMN_1854580	13	1.18E-09	-0.03	0.17
rs1477874	4	*PDGFRA* (intronic)	*SDHALP1*	ILMN_1734640	5	1.25E-09	-0.02	0.17
rs12240135	1	Upstream to *AX747377* and *NAV1*	*RAP1GDS1*	ILMN_1687724	4	1.34E-09	0.02	0.17
rs10207872	2	Intergenic	*NUSAP1*	ILMN_1726720	15	1.52E-09	-0.02	0.17
rs10873963	1	Upstream to *GIPC2*	*STAC2*	ILMN_1718295	17	1.55E-09	0.03	0.17
rs12965167	18	Intergenic	*RNF214*	ILMN_1800420	11	1.56E-09	-0.03	0.17
rs2716930	17	Intergenic	*ZNF550*	ILMN_1760102	19	1.69E-09	0.02	0.17
rs4140768	12	*ARHGDIB* (intronic)	*LOC647942*	ILMN_1705246	2	1.72E-09	-0.02	0.17
rs7598922	2	*DHX57* (missense)	*FLJ39632*	ILMN_1769704	14	1.93E-09	0.02	0.17

7.4 Discussion

7.4.1 Summary of results

In this analysis 25,269 *cis* acting and 5,898 *trans* acting eQTLs were identified within whole blood samples of 207 individuals. 4,829 of the *cis*-eQTLs (19%) were replications of eQTLs previously observed in a large analysis of blood eQTLs (Westra *et al*, 2013). As has been noted previously, in this study effect sizes for *cis*-eQTLs were larger as the distance between the SNP and gene probes decreases. However, in contrast to much of the existing literature (Petretto *et al*, 2006), *trans*-eQTL effect sizes were not smaller than effect sizes for *cis*-eQTLs. Using the identified *cis*-eQTLs, we observed enrichment amongst GWAS hits associated with any trait. One *cis*-eQTL locus was identified, at the gene *CERS5,* amongst those genetic variants which have shown suggestive association with antidepressant treatment response in a previous GWAS in the GENDEP sample. Annotation of the suggestive transcriptomic signatures predicting treatment response also revealed one gene probe (for *GSTM3*) under *cis*-regulation. Finally, a number of suggestive conditional *trans*-eQTLs were identified, where the relationship between genotype and gene expression level at baseline was related to the outcome of antidepressant treatment response.

7.4.2 Pattern of eQTLs

Of the *cis*-eQTLs identified here, 19% are exact replications (with matching SNP, gene probe and direction of effect) of *cis*-eQTLs previously identified in an analysis of whole blood samples taken from over 5,000 individuals (Westra *et al*, 2013). Although we do not replicate any *trans*-eQTL signals, this is likely to be due to methodological differences; the Westra dataset only considers SNPs included in the NHGRI catalogue of genome-wide significant variants (n=4,542), in contrast to our whole genome approach.

Indeed, much of the previous literature has focussed on *cis*-eQTLs with less research into the pattern of *trans*-eQTLs. This is in part due to the difficulties of the very large number of tests performed (in this analysis a total of 245,998,527,880 comparisons were performed). *Trans*-

eQTLs have not only been less frequently studied, but are also less reliably replicated between studies (Breitling *et al*, 2008). In light of this, we have used conservative corrections for multiple hypothesis testing; indeed many studies select thresholds separately for *cis*- and *trans*-eQTLs, whereas here we have corrected for all tests performed. But the lack of replication may also be due to both greater tissue- and context-specificity amongst *trans*-eQTLs (Fairfax *et al*, 2014; Price *et al*, 2011).

Despite these difficulties in assessing *trans*-eQTLs, they are of biological importance and we felt it was important to characterise them within this dataset. The heritability estimates of gene expression indicate that *trans*-eQTLs are likely to have an important contribution to variance in gene expression (Grundberg *et al*, 2012), and it is also thought that master regulators of gene expression (such as transcription factors) will be identified through the analysis of *trans*-eQTL hotspots, where a single genetic locus will exert effects on the transcript of a wide range of genes (Breitling *et al*, 2008). We did not identify any hotspots in this relatively small sample; the maximum number of gene probes associated with a single SNP amongst the *trans*-eQTLs was 12, and many of these genes were related ribosomal pseudogenes. Nevertheless, as more data become available, the presence of these potential hotspots can be better evaluated.

7.4.3 Using eQTLs as an annotation tool in GWAS

We observe significant enrichment of eQTL-SNPs amongst variants that have been previously linked to traits (consist with Nicolae *et al*, 2010), however the picture is less clear when focussing on those variants associated with antidepressant treatment response.

Taking those SNPs reaching suggestive significance in the GWAS of antidepressant treatment response previously conducted in the GENDEP sample (Uher *et al*, 2010), only one variant is located within a *cis*-eQTL locus. Within this locus, there are a cluster of 52 SNPs in high LD which all have *cis* effects on expression levels of the gene *CERS5*. The protein encoded by this gene is ceramide synthase 5; ceramide is a lipid molecule found in cell membranes which is involved in cellular signalling, such as cell differentiation and apoptosis (Obeid *et al*, 1993).

There are a number of known pathways for ceramide production, but ceramide has been implicated in a number of complex traits, included neuronal growth and development, with the synthesis and accumulation of ceramide known to be induced by stress stimuli (Bikman and Summers, 2011). Given the links between depression and both neuronal growth and stress, this might be an interesting candidate for further follow-up, albeit with caution as not only is this association unreplicated, but also recent work has highlighted how colocalising a genetic association with an eQTL is not sufficient prove that the eQTL is causally linked to the phenotype of interest (Giambartolomei *et al*, 2014; Plagnol *et al*, 2009).

As we have not found an enrichment of eQTLs amongst the suggestive genetic associations with response, it seems that using eQTL annotations to try to discriminate true signal from noise with antidepressant treatment response may be difficult to achieve given the current unreliability of eQTL annotation and the limited available sample sizes of antidepressant GWAS (where the signal to noise ratio may be too low). These problems have been also been highlighted as more general issues across a range of phenotypes when using *cis*-eQTL annotations to increase discovery for genome-wide analyses (Gagliano *et al*, 2014; Pickrell, 2014). This suggests that whilst eQTLs have potential as valuable annotation tools, further development is required to maximise their potential to gain insights into the biological pathways to disease.

7.4.4 Annotation of transcriptomic signals

When considering the potential transcriptomic predictors of treatment response, a *cis*-acting eQTL locus controlling expression levels of *GSTM3* was identified. This gene appears to be of relevance to antidepressant response given its expression in the brain, and its role in oxidative stress, drug uptake and detoxification. Furthermore, in a chronic restraint model of depression, changes in expression levels of *GSTM3*, amongst other genes, were observed in rat brain, which were reversed with antidepressant treatment (Jungke *et al*, 2011). Whilst the association between gene expression levels prior to treatment and treatment response is only suggestive (p<0.0001), the link between gene expression differences and genotypic variability point to potential genotypic markers of treatment response.

7.4.5 Treatment response conditional eQTLs

Finally, exploring the relationship between genetic control of gene expression and the phenotype of antidepressant treatment response, no significant response-conditional eQTLs were observed. However, a number of suggestive *trans*-eQTLs were found. This links with previous indications that *trans*-eQTLs are more context-specific than *cis*-eQTLs (Fairfax *et al*, 2014; van Nas *et al*, 2010), although as noted above, the pattern may be an artefact of poor reliability.

The response conditional eQTLs may indicate that the genetic control of gene expression varies in a way which is predictive of later treatment response, and point to potential biological differences between those who are more or less likely to respond to antidepressant treatment. Therefore the SNPs and genes implicated in this analysis may point to the molecular pathways involved in antidepressant response. However, annotation within the suggestive conditional *trans*-eQTL probes identified here did not reveal any previously described pathways.

7.4.6 Tissue-specificity of eQTLs

All eQTL associations must be interpreted within the context of the observed tissue-specificity of eQTLs (as discussed in the Introduction). This analysis has been performed using peripheral blood samples and ease of collection means that this is usually the case. However, the primary tissue of interest for antidepressant response is the brain. It may be that there are phenotype-relevant eQTLs that we are unable to detect in this analysis as they are only seen in brain tissue (or vice versa, some of the signals detected here may not be relevant as they are not present at the site of antidepressant action). Nevertheless, obtaining large samples of brain tissue for transcriptomic analysis is complex, with associated difficulties of post mortem changes in expression profiles and mRNA quality to be considered. Additionally, brain region-specific eQTLs have been reported (Hernandez *et al*, 2012), and the site of action for antidepressants is not known. Finally, to consider conditional eQTLs (where genotypic control of gene expression

prior to treatment is considered in the context of response after receiving treatment) would not be possible using brain tissue.

7.4.7 Conclusions

This chapter explores the pattern of eQTLs observed within the GENDEP dataset. Using these eQTLs, it has been possible to identify a *cis*-eQTL locus which controls expression levels of the gene *CERS5* amongst the suggestive variants implicated in a previous GWAS on the GENDEP sample, as well as a *cis*-eQTL locus for the gene *GSTM3* amongst the suggestive transcripts implicated in a previous transcriptomic analysis of the GENDEP sample.

Both of these candidates show promise in terms of the relevance of their known functions to antidepressant treatment outcomes, and replication of the identified eQTLs are observed in other datasets. However, it should be remembered that evidence linking these genes to treatment response remains suggestive and unreplicated at this stage, and evidence causally linking the eQTL to the phenotype has not be found.

Nonetheless, the analysis presented here demonstrates how understanding the genetic effects on gene expression levels may be useful for identifying trait-associated genetic variants, and understanding how they exert their effects. Furthermore, the approach highlights that if phenotype-dependent eQTLs can be identified, they may give insight into the biological differences underlying that phenotype. The potential value of eQTLs for dissecting molecular consequences of genetic variability will only increase in the future, as our understanding and characterisation of eQTLs improves.

Chapter 8 Discussion

8.1 Summary of principal findings

In this thesis I aimed to identify genetic biomarkers of antidepressant outcomes, within the GENDEP dataset. I have used a number of different approaches, including candidate gene, transcriptomic and eQTL-based methodologies to explore biomarkers that predict treatment outcomes (both wanted and unwanted) and index the biological mechanisms that underpin therapeutic efficacy.

8.1.1 Predicting treatment outcomes

8.1.1.1 Genetic predictors of antidepressant side effects

After grouping the many different antidepressant side effects using knowledge about their pharmacological basis, candidate gene analysis revealed that risk of serotonergic side effects was associated with variation in the *HTR2C* gene. However, when attempting replication of these findings in GenPod (a pharmacogenetic study of antidepressants with similar design features to GENDEP), this association was not observed. Nevertheless, this approach demonstrates that the rational grouping of side effects based on their known pharmacological causes may be a useful method to condense the large number of different side effects reported with antidepressant treatment.

As Kato and Serretti (2010) highlight, the large number of different possible outcomes has been a key limitation in the current literature regarding the genetic predictors of these ADRs. Therefore the work presented in Chapter 3 offers a novel method which may enable researchers to gain traction on this issue, condensing the large number of reported outcomes into a smaller number of variables, without losing the detail necessary to capture the important mechanistic differences between side effects. This is the key message of the analysis; nonetheless we have been able to highlight genetic associations with antidepressant side effects which align with existing literature on the role of HTR2C in side effects associated with antipsychotic treatment (Reynolds et al, 2005), although the failure to replicate this association does mean that further investigation of the findings is necessary before conclusions can be drawn.

Whilst the candidate gene approach used in Chapter 3 has the limitation of not surveying the entire genome, there is strong pharmacological evidence of the mechanisms underlying side effects which justifies the selection of the candidates used. Further, given the profiles of ADRS across the course of the 12 week study, a longitudinal model is an important strength of the approach which could not be easily employed within a genome-wide approach. Nevertheless, the limitations of candidate gene studies are discussed within the introduction, and the important of replication and validation of these analyses is key.

8.1.1.2 Drug metabolism variables in treatment outcomes

When exploring the pharmacokinetics of antidepressant response, genotypic variability in the cytochrome P450 enzymes was found to be unrelated to treatment response, side effects or study discontinuation. Furthermore, serum concentrations of antidepressants were not predictive of either response or the majority of side effects measured. Associations between serum drug concentration and dry mouth were observed for patients taking nortriptyline, and for patients taking escitalopram, drug concentration was linked to the rarer side effects of diarrhoea and dizziness. Nevertheless, beyond these specific ADRs, the findings here show that the observed variability in treatment outcomes is not due to pharmacokinetic factors.

The data presented in Chapter 4 builds on previous work linking CYP450 genotypes to serum levels of antidepressant (for example, Rudberg *et al*, 2006; Rudberg *et al*, 2008; Kirchheiner *et al*, 2004), giving further information on the relationship within a large, well-phenotyped patient sample receiving escitalopram or nortriptyline.

When addressing the key clinical question of whether these effects on rates of drug metabolism have an impact on treatment outcomes, the previously published, extensive review from Matchar *et al* highlighted the limitations in the available evidence in 2007, with a paucity of studies examining the link between CP450 genotype and treatment outcomes. The evidence presenting in this thesis, together with other publications published since 2007 (including Peters

et al 2008; Mrazek *et al* 2011; Serretti *et al* 2009) all represent important efforts to tackle this issue. Generally, our findings align with these other reports, even in papers where associations are reported (for example Mrazek *et al* 2011) these effects are generally small in size, and do not survive correction for multiple hypothesis testing, suggesting that they lack the robustness required to be translationally valuable.

The absence of association between CYP450 genotype and treatment outcomes in clinical cohorts is not surprising, given the therapeutic windows of antidepressants. Unlike drugs such as warfarin, where the therapeutic window is very small due to the severe and lethal side effects associated with high levels of drug, many antidepressants (in particular SSRIs) have relatively large therapeutic windows, with low risk of severe adverse drug reactions when higher doses are prescribed. Therefore, clinicians are able to adjust the prescribed dose in response to poor efficacy or side effects.

It is important to remember that the design of GENDEP means that it is only possible to conclude the extent to which CYP450 genotypes are linked to treatment outcomes in patients under clinical observation, where doses can be adjusted in response to side effects or lack of efficacy. Nevertheless, the fixed-dosage study published by Steimer *et al* (2005) looking at patients taking amitriptyline also failed to find an association with treatment response. Putting the evidence presented within this thesis together with the other literature published in the area, it would seem that the potential value of CYP450 genotyping of patients who are to be prescribed antidepressants is far from being established, and should not be recommended to clinicians without substantial further evidence, and clear demonstration that treatment guided by CYP450 genotyping can improve patient outcomes.

8.1.2 Understanding the biological mechanisms underpinning response

8.1.2.1 Gene expression correlates of treatment response

Analysis of transcriptomic changes associated with antidepressant treatment indicated that response-correlated changes in gene expression levels occur at a network level. One module of coexpressed genes showed significant correlation with treatment response, and this relationship

appeared to be general rather than drug-specific. Annotation of the implicated module highlights pathways related to “Cellular Assembly and Organisation, Nervous System Development and Function, Cell Signalling”, which is consistent with existing neurotrophic theories of antidepressant action. Therefore, the results indicate that systems-level investigations into neurotrophic processes may prove fruitful in unveiling the mechanisms of antidepressant action.

As the largest analysis of transcriptomic changes associated with antidepressant treatment to date, these findings are useful for placing previous research looking at expression changes within individual candidate genes in context. The findings presented here indicate that there are no single genes which show large changes in expression level, instead complex network-level changes are seen to be associated with treatment response. This evidence is consistent not only with what is known about the effects of antidepressants on neurobiology but also with the accumulating literature demonstrating that DNA variation in single genes do not have a predictive effect on treatment outcomes (Tansey *et al*, 2013), in that it indicates that antidepressant treatment response is a complex phenotype with multiple factors interacting in a complex system. Nevertheless, the data presented here is still from a limited sample size; to identify which genes are involved in these networks underlying treatment response, larger samples are needed, and replication of these findings is required. In particular, it is important that these findings are replicated with reference to treatment response outcomes (as opposed to simply taking antidepressant medications), as these will be the most informative with respect to understanding the mechanisms of effective antidepressant treatment.

8.1.2.2 Identification and exploration of eQTLs

Moving to more generally consider the mechanism through which genetic differences might influence gene expression, the pattern of eQTLs within this dataset was established, with 25,269 *cis*- and 5,898 *trans*-eQTLs observed. Exploring the value of this data as a potential bioinformatics tool, eQTL annotations could be made for two previously identified genetic (*CERS5*) and transcriptomic (*GSTM3*) predictors of treatment response, highlighting them as possible candidates for further follow-up.

Whilst the work presented here has been conducted on a sample size that is much smaller than many other studies exploring eQTLs (for example, Westra *et al* (2013) included data from 5,311 individuals in their meta-analysis with an additional 2,775 in a replication set), it is encouraging to note that within GENDEP it is possible to replicate many of the eQTLs that have previously been identified.

The key purpose of Chapter 7 was to demonstrate how the identification of eQTLs can be exploited to identify important candidates within suggestive findings for further follow up and to begin to build a pathway between genotype and phenotype. To this end, the results are intriguing. Whilst the evidence linking *CERS5* and *GSTM3* to treatment response in GENDEP is only suggestive, the annotation of these candidates as eQTLs in the analysis presented here is also observed within other datasets. Furthermore, the known action of these genes links very well with what is known about the pathways underpinning antidepressant efficacy. *CERS5* encodes ceramide synthase 5; ceramide is not only linked to neuronal development but its synthesis is linked to stress-related stimuli (Bikman and Summer 2011). Similarly, *GSTM3* encodes glutathione S-transferase, which is expressed in the brain, plays a key role in drug uptake and has previously been linked to depression and antidepressant phenotypes in animal models (Jungke *et al* 2011). Nevertheless, the important question of the tissue specificity of these eQTLs is yet to be addressed, and identification of an eQTL at a gene of interest does not prove that this eQTL is causally linked to the phenotype of interest.

Finally, the evidence within this chapter also includes a preliminary investigation of how the context-specificity of eQTLs demonstrated by Fairfax *et al* (2014) might also be observed when considering clinical phenotypes. Whilst no antidepressant-response conditional eQTLs reached significance, a number of suggestive conditional trans-eQTLs were observed. Further exploration of these conditional eQTLs is needed to confirm whether these observations are replicable.

8.2 Limitations

Whilst the limitations specific to each analysis have been outlined within the relevant chapters, there are some overarching limitations that are important to bear in mind when interpreting the evidence presented here.

One issue which is recurrent throughout this thesis is that of statistical power. GENDEP is amongst the largest studies considering the genetics of antidepressant treatment response, alongside STAR*D (Rush *et al*, 2004), GenPod (Thomas *et al*, 2008) and MARS (Hennings *et al*, 2009). But whilst these datasets have been critical in establishing the genetic architecture of treatment response as a complex polygenic trait, even larger samples will be required to identify genetic biomarkers of treatment outcomes.

In order to obtain a sample of 868 patients, GENDEP incorporated several pragmatic design features to broaden recruitment potential. These included non-random allocation to treatment in cases where patients had existing contra-indications to one of the study medications, the absence of a placebo arm, and a flexible dosing protocol rather than a fixed dosage design. Each of these features may add confounds to the design, but as March *et al.* (2005) have highlighted, the inclusion of more patients through these pragmatic features increases the generalizability of findings. Nevertheless, the requirement that all patients were Caucasian (to limit population stratification in genetic analyses), as well as the weekly clinical contact (either via telephone or face-to-face interview), may be of importance when considering translation into the clinic.

A further consideration is that GENDEP focusses on only two of the many antidepressants currently available. The selection of escitalopram and nortriptyline as study medications was designed to give maximum value, as these drugs have similar efficacy but differentially affect two of the key neurotransmitter systems targeted by antidepressants. Nevertheless, the relevance of findings within this sample to patients taking other medications must be established. This is particularly the case with antidepressants which have mechanisms of action

involving different neurotransmitter systems (for example, agomelatine which targets melatonin and serotonin 2C receptors).

Finally, a key issue to consider in any research effort exploring antidepressant treatment is the intense debate surrounding the issue of antidepressant efficacy relative to placebo (Horder *et al*, 2011; Kirsch *et al*, 2008), the impact of publication bias in the literature (Turner *et al*, 2008) and the relative efficacy of different medications (Cipriani *et al*, 2009; Gartlehner *et al*, 2011). This debate is not limited to the use of medications for the treatment of MDD, but also applies to the evidence regarding psychotherapy (for example, discussed by Nutt and Sharpe, 2008). Nonetheless, the only way to tackle the debate about the benefits of these medications is through further research, to better understand how they exert their effects, and why outcomes are so variable between patients.

8.3 Implications of findings

8.3.1 Clinical implications

Given the absence of any replicated genetic biomarkers of treatment outcomes, the research presented here is not at a stage where findings can be translated into the clinic. Nevertheless, given the interest regarding the potential for pharmacogenetics in guiding treatment, effective communication of findings to clinicians is important. This is particularly relevant for the null findings regarding the role of pharmacokinetics on variability in outcomes (presented in Chapter 4 & 5), given the hype surrounding the potential of CYP450 genotyping (as discussed by Matchar *et al*, 2007a). Whilst the study design of GENDEP means that it is not possible to draw conclusions regarding the value of genotyping prior to treatment, on reviewing the literature in this field, there is also a scarcity of well-powered studies to address this question. To ensure that clinicians can base their treatment recommendations on the best available evidence, it is important to ensure that null results receive equal attention to positive findings.

8.3.2 Research implications

Whilst the findings presented here cannot yet be used to guide treatment or the development of novel medications, they do have broader implications for further research efforts. Both within this thesis and more broadly across pharmacogenetics of antidepressant response, findings indicate that treatment response is a complex phenotype almost certainly involving many genes of small effect (GENDEP Investigators; MARS Investigators; STAR*D Investigators, 2013; Tansey *et al*, 2013; Tansey *et al*, 2012). This is counter to the early hopes that pharmacogenetic phenotypes might prove simpler than common familial disease phenotypes, but fits with the complexity and redundancy that we know exists within the brain. Furthermore, given that depression is a complex phenotype likely to involve many different molecular components, then it is fitting that effective treatment mirrors this complexity.

The evidence regarding antidepressant-linked side effects suggests a similar complex genetic basis, but in light of the importance of ADRs for treatment outcomes, similar efforts to empirically establish the role of genetics in ADRs using a genome-wide complex trait analysis (GCTA) approach would be valuable. However, the additional complexity of many different outcomes has limited research to date.

In the context of this complex polygenic phenotype, “omic” approaches are needed, where the entire genome or transcriptome is considered. This is clearly demonstrated by the transcriptomic analysis in Chapter 6, where correlates of treatment response were only identified using system level analyses. This was the case because the signal of changes in gene expression levels was distributed across many genes within the same network.

Furthermore, in the face of this phenotypic complexity, this thesis has demonstrated the value of layering together different levels of information to dissect the molecular mechanisms involved in antidepressant action. This involves not only using pharmacological information (as in Chapters 3, 4 and 5), but also the use of different layers of genetic data (as shown in Chapter 7). Further

approaches considering methylomic, proteomic and metabolomic data may also prove fruitful in tracing the molecular pathway of antidepressant action.

8.4 Future directions

Looking ahead to how research in this field may progress, there are two complementary approaches that will be valuable in driving forward our understanding of the genetic underpinnings of antidepressant action.

The first approach is to continue with the development of more sophisticated methodological approaches to genome-wide analysis. This may include using bioinformatic data, such as the application of functional annotations (explored with the eQTL analysis presented in Chapter 7). It has been shown that within a Bayesian framework, it is possible to extract more signal from genomic analyses (Gagliano *et al*, 2014), and with improved annotation data, the success of the approach should increase further. Additionally, given the polygenic nature of the phenotype, multivariate methods which allow exploration of how effects might be distributed across many genetic predictors are likely to be of value. For example, machine learning techniques can be used to capture a signal which is diffusely spread across a large number of predictor variables, and are increasingly being considered within the field of genetics, exploring both genomic (Bureau *et al*, 2005) and transcriptomic (Pirooznia *et al*, 2008) data.

The second approach is to gain more statistical power for genome-wide analyses by increasing sample sizes, using strategies that will enable easier and cheaper collection of genetic and treatment response data from patients. One potential avenue by which this could be achieved is through the use of electronic medical records. Schemes such as the NIHR Bioresource (http://bioresource.nihr.ac.uk/), involve large scale collection of biological samples for genomic analyses, which are then linked to the information that is contained within patient records. These datasets can then be mined for patients with relevant phenotypic features. Given the high prescription rates for antidepressants, these methods could potentially collect information on very large numbers of patients, albeit at the cost of increased sample heterogeneity, and the

non-random allocation of patients to treatment. Furthermore, developments in mobile technology and online assessment tools allow the easy administration of questionnaires to rate symptoms and side effects. If data from these sources could be integrated into medical records, then in-depth, longitudinal data on treatment outcomes would be available without the need for researchers to undertake face-to-face or telephone assessments for all study participants.

8.5 Conclusions

Major depressive disorder is a huge public health issue, and is projected to be the number one cause of disease burden across the globe by 2030 (World Health Organisation, 2008). But currently, our ability to treat the illness is limited, with only 30% of patients achieving remission with their first treatment medication (Trivedi *et al*, 2006). However, the variability that is seen in treatment outcomes is known to be determined, in part, by genetics (Tansey *et al*, 2013). It is not yet possible to identify the specific genes that are involved in treatment efficacy, but examples of genetic research in other traits such as diabetes (Morris *et al*, 2012) or schizophrenia (Schizophrenia Working Group of the Psychiatric Genomics, 2014) indicate that continued efforts to collect larger samples will be successful. Once genetic biomarkers have been identified, they have huge potential in improving treatment options for MDD. The identification of response-linked genes gives an important window into the biology of antidepressant action, which may enable the development of novel treatments. Furthermore, if these markers can be used to predict outcomes prior to treatment, they have the potential to shorten treatment times, improve response rates and reduce side effects, with only a simple and relatively low cost genetic test required. Therefore, the continued research effort into identifying genetic predictors of antidepressant outcomes has the potential to have a real impact on the treatment options available to patients with depression.

Chapter 9 References

Abbas AR, Wolslegel K, Seshasayee D, Modrusan Z, Clark HF (2009). Deconvolution of blood microarray data identifies cellular activation patterns in systemic lupus erythematosus. *PLoS One* **4**(7): e6098.

Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA*, et al* (2010). A map of human genome variation from population-scale sequencing. *Nature* **467**(7319): 1061-1073.

Adkins DE, Clark SL, Aberg K, Hettema JM, Bukszar J, McClay JL*, et al* (2012). Genome-wide pharmacogenomic study of citalopram-induced side effects in STAR*D. *Transl Psychiatry* **2**: e129.

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn. American Psychiatric Association: Washington DC.

Asberg M, Cronholm B, Sjoqvist F, Tuck D (1971). Relationship between plasma level and therapeutic effect of nortriptyline. *British medical journal* **3**(5770): 331-334.

Bagby RM, Ryder AG, Schuller DR, Marshall MB (2004). The Hamilton Depression Rating Scale: has the gold standard become a lead weight? *Am J Psychiatry* **161**(12): 2163-2177.

Barrett JC, Fry B, Maller J, Daly MJ (2005). Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21(2): 263-265.

Baumann P (1996). Pharmacokinetic-pharmacodynamic relationship of the selective serotonin reuptake inhibitors. *Clin Pharmacokinet* **31**(6): 444-469.

Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J (1961). An inventory for measuring depression. *Arch Gen Psychiatry* **4**: 561-571.

Belzeaux R, Bergon A, Jeanjean V, Loriod B, Formisano-TrÃ©ziny C, Verrier L*, et al* (2012). Responder and nonresponder patients exhibit different peripheral transcriptional signatures during major depressive episode. *Translational Psychiatry* **2**(11).

Belzeaux R, Formisano-Treziny C, Loundou A, Boyer L, Gabert J, Samuelian JC*, et al* (2010). Clinical variations modulate patterns of gene expression and define blood biomarkers in major depression. *Journal of psychiatric research* **44**(16): 1205-1213.

Benjamini Y, Hochberg Y (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B (Methodological)* **57**(1): 289-300.

Berton O, Nestler EJ (2006). New approaches to antidepressant drug discovery: beyond monoamines. *Nat Rev Neurosci* **7**(2): 137-151.

Bijl MJ, Visser LE, Hofman A, Vulto AG, Van Gelder T, Stricker BHC*, et al* (2008). Influence of the CYP2D6*4 polymorphism on dose, switching and discontinuation of antidepressants. *British Journal of Clinical Pharmacology* **65**(4): 558-564.

Bikman BT, Summers SA (2011). Ceramides as modulators of cellular and whole-body metabolism. The Journal of Clinical Investigation 121(11): 4222-4230.

Biswas S, Storey J, Akey J (2008). Mapping gene expression quantitative trait loci by singular value decomposition and independent component analysis. *BMC Bioinformatics* **9**(1): 244.

Boothman LJ, Mitchell SN, Sharp T (2006). Investigation of the SSRI augmentation properties of 5-HT(2) receptor antagonists using in vivo microdialysis. *Neuropharmacology* **50**(6): 726-732.

Bostwick JM (2010). A generalist's guide to treating patients with depression with an emphasis on using side effects to tailor antidepressant therapy. *Mayo Clin Proc* **85**(6): 538-550.

Brandl EJ, Tiwari AK, Zhou X, Deluce J, Kennedy JL, Muller DJ*, et al* (2014). Influence of CYP2D6 and CYP2C19 gene variants on antidepressant response in obsessive-compulsive disorder. *Pharmacogenomics J* **14**(2): 176-181.

Breitling R, Li Y, Tesson BM, Fu J, Wu C, Wiltshire T*, et al* (2008). Genetical genomics: spotlight on QTL hotspots. *PLoS Genet* **4**(10): e1000232.

Bull SA, Hunkeler EM, Lee JY, Rowland CR, Williamson TE, Schwab JR*, et al* (2002). Discontinuing or switching selective serotonin-reuptake inhibitors. *Ann Pharmacother* **36**(4): 578-584.

Bureau A, Dupuis J, Falls K, Lunetta KL, Hayward B, Keith TP*, et al* (2005). Identifying SNPs predictive of phenotype using random forests. *Genetic epidemiology* **28**(2): 171-182.

Cai C, Langfelder P, Fuller TF, Oldham MC, Luo R, van den Berg LH*, et al* (2010). Is human blood a good surrogate for brain tissue in transcriptional studies? *BMC Genomics* **11**: 589.

Carruthers SG (1994). Adverse effects of alpha 1-adrenergic blocking drugs. *Drug Saf* **11**(1): 12-20.

Castrén E, Võikar V, Rantamäki T (2007). Role of neurotrophic factors in depression. *Current opinion in pharmacology* **7**(1): 18-21.

Cattaneo A, Bocchio-Chiavetto L, Zanardini R, Milanesi E, Placentino A, Gennarelli M (2010). Reduced peripheral brain-derived neurotrophic factor mRNA levels are normalized by antidepressant treatment. *Int J Neuropsychopharmacol* **13**(1): 103-108.

Cattaneo A, Gennarelli M, Uher R, Breen G, Farmer A, Aitchison KJ*, et al* (2013). Candidate genes expression profile associated with antidepressants response in the GENDEP study: differentiating between baseline 'predictors' and longitudinal 'targets'. *Neuropsychopharmacology* **38**(3): 377-385.

Chen S, Chou WH, Blouin RA, Mao Z, Humphries LL, Meek QC*, et al* (1996). The cytochrome P450 2D6 (CYP2D6) enzyme polymorphism: screening costs and influence on clinical outcomes in psychiatry. *Clin Pharmacol Ther* **60**(5): 522-534.

Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JPT, Churchill R*, et al* (2009). Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *The Lancet* **373**(9665): 746-758.

Clark SL, Adkins DE, Aberg K, Hettema JM, McClay JL, Souza RP*, et al* (2012). Pharmacogenomic study of side-effects for antidepressant treatment options in STAR*D. *Psychol Med* **42**(6): 1151-1162.

Conney AH, Buening MK, Pantuck EJ, Pantuck CB, Fortner JG, Anderson KE*, et al* (1980). Regulation of human drug metabolism by dietary factors. *Ciba Foundation symposium* **76**: 147-167.

Couchman L (2012). Turbulent flow chromatography in bioanalysis: a review. *Biomedical chromatography : BMC* **26**(8): 892-905.

Crawford AA, Lewis S, Hodgson K, McGuffin P, Aitchison KJ, Nutt D*, et al* (in prep). Are genetic variants able to differentially predict sexual dysfunction in men following treatment with citalopram or reboxetine? Findings from a randomised control study.

Cremers TIFH, Giorgetti M, Bosker FJ, Hogg S, Arnt J, Mork A*, et al* (2004). Inactivation of 5-HT2C Receptors Potentiates Consequences of Serotonin Reuptake Blockade. *Neuropsychopharmacology* **29**(10): 1782-1789.

Dahl ML, Bertilsson L, Nordin C (1996). Steady-state plasma levels of nortriptyline and its 10-hydroxy metabolite: relationship to the CYP2D6 genotype. *Psychopharmacology* **123**(4): 315-319.

de Jong S, Boks MPM, Fuller TF, Strengman E, Janson E, de Kovel CGF*, et al* (2012). A Gene Co-Expression Network in Whole Blood of Schizophrenia Patients Is Independent of Antipsychotic-Use and Enriched for Brain-Expressed Genes. *PLoS ONE* **7**(6): e39498.

de Leon J, Susce MT, Murray-Carmichael E (2006). The AmpliChip CYP450 genotyping test: Integrating a new clinical tool. *Molecular diagnosis & therapy* **10**(3): 135-151.

de Morais SM, Wilkinson GR, Blaisdell J, Nakamura K, Meyer UA, Goldstein JA (1994). The major genetic defect responsible for the polymorphism of S-mephenytoin metabolism in humans. *The Journal of biological chemistry* **269**(22): 15419-15422.

Degner JF, Pai AA, Pique-Regi R, Veyrieras J-B, Gaffney DJ, Pickrell JK*, et al* (2012). DNase[thinsp]I sensitivity QTLs are a major determinant of human expression variation. *Nature* **482**(7385): 390-394.

Dimas AS, Deutsch S, Stranger BE, Montgomery SB, Borel C, Attar-Cohen H*, et al* (2009). Common Regulatory Variation Impacts Gene Expression in a Cell Type–Dependent Manner. *Science* **325**(5945): 1246-1250.

Dimas AS, Nica AC, Montgomery SB, Stranger BE, Raj T, Buil A*, et al* (2012). Sex-biased genetic effects on gene regulation in humans. *Genome research* **22**(12): 2368-2375.

Dowlati Y, Herrmann N, Swardfager W, Liu H, Sham L, Reim EK*, et al* (2010). A meta-analysis of cytokines in major depression. *Biol Psychiatry* **67**(5): 446-457.

Dufour H, Bouchacourt M, Thermoz P, Viala A, Phak Rop P, Gouezo F*, et al* (1987). Citalopram--a highly selective 5-HT uptake inhibitor--in the treatment of depressed patients. *Int Clin Psychopharmacol* **2**(3): 225-237.

Duman RS, Heninger GR, Nestler EJ (1997). A molecular and cellular theory of depression. *Archives of General Psychiatry* **54**(7): 597-606.

Dunning M, Lynch A, Eldridge M illuminaHumanv4.db: Illumina HumanHT12v4 annotation data (chip illuminaHumanv4). R package version 1.22.1.

EGAPP Working Group (2007). Recommendations from the EGAPP Working Group: testing for cytochrome P450 polymorphisms in adults with nonpsychotic depression treated with selective serotonin reuptake inhibitors. *Genet Med* **9**(12): 819-825.

Eichelbaum M, Spannbrucker N, Steincke B, Dengler HJ (1979). Defective N-oxidation of sparteine in man: a new pharmacogenetic defect. *Eur J Clin Pharmacol* **16**(3): 183-187.

Eyding D, Lelgemann M, Grouven U, Härter M, Kromp M, Kaiser T*, et al* (2010). Reboxetine for acute treatment of major depression: systematic review and meta-analysis of published and unpublished placebo and selective serotonin reuptake inhibitor controlled trials. *BMJ* **341**.

Fairfax BP, Humburg P, Makino S, Naranbhai V, Wong D, Lau E*, et al* (2014). Innate Immune Activity Conditions the Effect of Regulatory Variants upon Monocyte Gene Expression. *Science* **343**(6175).

Faul F, Erdfelder E, Lang AG, Buchner A (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* **39**(2): 175-191.

FDA (2001). Guidance for industry. Bioanalytical method validation.

FDA (2011). FDA: Drug Interactions (CYP450 inhibitors and inducers). Vol 2013.

Fehrmann RSN, Jansen RC, Veldink JH, Westra H-J, Arends D, Bonder MJ*, et al* (2011). Trans-eQTLs Reveal That Independent Genetic Variants Associated with a Complex Phenotype Converge on Intermediate Genes, with a Major Role for the HLA. *PLoS Genet* **7**(8): e1002197.

Ferguson RJ, De Morais SM, Benhamou S, Bouchardy C, Blaisdell J, Ibeanu G*, et al* (1998). A new genetic defect in human CYP2C19: mutation of the initiation codon is responsible for poor metabolism of S-mephenytoin. *The Journal of pharmacology and experimental therapeutics* **284**(1): 356-361.

Foglia JP, Pollock BG, Kirshner MA, Rosen J, Sweet R, Mulsant B (1997). Plasma levels of citalopram enantiomers and metabolites in elderly patients. *Psychopharmacology bulletin* **33**(1): 109-112.

Franchini L, Serretti A, Gasperini M, Smeraldi E (1998). Familial concordance of fluvoxamine response as a tool for differentiating mood disorder pedigrees. *Journal of Psychiatric Research* **32**(5): 255-259.

Frank E, Prien RF, Jarrett RB, et al. (1991). Conceptualization and rationale for consensus definitions of terms in major depressive disorder: Remission, recovery, relapse, and recurrence. *Archives of General Psychiatry* **48**(9): 851-855.

Frazer A, Benmansour S (2002). Delayed pharmacological effects of antidepressants. *Mol Psychiatry* **7 Suppl 1**: S23-28.

Freeman B, Smith N, Curtis C, Huckett L, Mill J, Craig IW (2003). DNA from buccal swabs recruited by mail: evaluation of storage effects on long-term stability and suitability for multiplex polymerase chain reaction genotyping. *Behav Genet* **33**(1): 67-72.

Fujitsuka N, Asakawa A, Hayashi M, Sameshima M, Amitani H, Kojima S*, et al* (2009). Selective Serotonin Reuptake Inhibitors Modify Physiological Gastrointestinal Motor Activities via 5-HT2c Receptor and Acyl Ghrelin. *Biological Psychiatry* **65**(9): 748-759.

Gaffney DJ, Veyrieras JB, Degner JF, Pique-Regi R, Pai AA, Crawford GE*, et al* (2012). Dissecting the regulatory architecture of gene expression QTLs. *Genome biology* **13**(1): R7.

Gage BF, Eby C, Johnson JA, Deych E, Rieder MJ, Ridker PM*, et al* (2008). Use of pharmacogenetic and clinical factors to predict the therapeutic dose of warfarin. *Clin Pharmacol Ther* **84**(3): 326-331.

Gagliano SA, Barnes MR, Weale ME, Knight J (2014). A Bayesian Method to Incorporate Hundreds of Functional Characteristics with Association Evidence to Improve Variant Prioritization. *PLoS ONE* **9**(5): e98122.

Garriock HA, Kraft JB, Shyn SI, Peters EJ, Yokoyama JS, Jenkins GD*, et al* (2010). A genomewide association study of citalopram response in major depressive disorder. *Biol Psychiatry* **67**(2): 133-138.

Gartlehner G, Hansen RA, Morgan LC, Thaler K, Lux L, Van Noord M*, et al* (2011). Comparative benefits and harms of second-generation antidepressants for treating major depressive disorder: an updated meta-analysis. *Annals of internal medicine* **155**(11): 772-785.

Gaujoux R, Seoighe C (2013). CellMix: a comprehensive toolbox for gene expression deconvolution. *Bioinformatics* **29**(17): 2211-2212.

GENDEP Investigators; MARS Investigators; STAR*D Investigators (2013). Common genetic variation and antidepressant efficacy in major depressive disorder: a meta-analysis of three genome-wide pharmacogenetic studies. *Am J Psychiatry* **170**(2): 207-217.

Gerhold DL, Jensen RV, Gullans SR (2002). Better therapeutics through microarrays. *Nat Genet* **32 Suppl**: 547-551.

Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C*, et al* (2014). Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet* **10**(5): e1004383.

Gibbs JR, van der Brug MP, Hernandez DG, Traynor BJ, Nalls MA, Lai S-L*, et al* (2010). Abundant Quantitative Trait Loci Exist for DNA Methylation and Gene Expression in Human Brain. *PLoS Genet* **6**(5): e1000952.

Gillman PK (2007). Tricyclic antidepressant pharmacology and therapeutic drug interactions updated. *British journal of pharmacology* **151**(6): 737-748.

Glass D, Vinuela A, Davies MN, Ramasamy A, Parts L, Knowles D*, et al* (2013). Gene expression changes with age in skin, adipose tissue, blood and brain. *Genome biology* **14**(7): R75.

Grayson D, Bridges K, Cook D, Goldberg D (1990). The validity of diagnostic systems for common mental disorders: a comparison between the ID-CATEGO and the DSM-III systems. *Psychol Med* **20**(1): 209-218.

Greenberg RP, Bornstein RF, Greenberg MD, Fisher S (1992). A meta-analysis of antidepressant outcome under "blinder" conditions. *Journal of consulting and clinical psychology* **60**(5): 664-669; discussion 670-667.

Grundberg E, Small KS, Hedman AK, Nica AC, Buil A, Keildson S*, et al* (2012). Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nat Genet* **44**(10): 1084-1089.

GTEx Consortium (2013). The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45**(6): 580-585.

Guengerich FP (2008). Cytochrome p450 and chemical toxicology. *Chemical research in toxicology* **21**(1): 70-83.

Gutierrez M, Mengel H (2002). Pharmacokinetics of escitalopram. *National Institute of Mental Health 42nd Annual New Clinical Drug Evaluation Unit Meeting*: Boca Raton, FL, USA.

Gutierrez MM, Rosenberg J, Abramowitz W (2003). An evaluation of the potential for pharmacokinetic interaction between escitalopram and the cytochrome P450 3A4 inhibitor ritonavir. *Clinical Therapeutics* **25**(4): 1200-1210.

Hall-Flavin DK, Winner JG, Allen JD, Jordan JJ, Nesheim RS, Snyder KA*, et al* (2012). Using a pharmacogenomic algorithm to guide the treatment of depression. *Transl Psychiatry* **2**: e172.

Hamilton M (1960). A rating scale for depression. *J Neurol Neurosurg Psychiatry* **23**: 56-62.

Hamilton M (1967). Development of a rating scale for primary depressive illness. *Br J Soc Clin Psychol* **6**(4): 278-296.

Hammer W, Sjoqvist F (1967). Plasma levels of monomethylated tricyclic antidepressants during treatment with imipramine-like compounds. *Life sciences* **6**(17): 1895-1903.

Hamon M, Bourgoin S (2006). Pharmacological profile of antidepressants: a likely basis for their efficacy and side effects? *European Neuropsychopharmacology* **16**(Supplement 5): S625-S632.

Harmer CJ, Cowen PJ (2013). 'It's the way that you look at it'—a cognitive neuropsychological account of SSRI action in depression. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368**(1615).

Heils A, Teufel A, Petri S, Stober G, Riederer P, Bengel D*, et al* (1996). Allelic variation of human serotonin transporter gene expression. *J Neurochem* **66**(6): 2621-2624.

Hennings JM, Owashi T, Binder EB, Horstmann S, Menke A, Kloiber S*, et al* (2009). Clinical characteristics and treatment outcome in a representative sample of depressed inpatients - findings from the Munich Antidepressant Response Signature (MARS) project. *J Psychiatr Res* **43**(3): 215-229.

Hernandez DG, Nalls MA, Moore M, Chong S, Dillman A, Trabzuni D*, et al* (2012). Integration of GWAS SNPs and tissue specific expression profiling reveal discrete eQTLs for human traits in blood and brain. *Neurobiol Dis* **47**(1): 20-28.

Hicks JK, Swen JJ, Thorn CF, Sangkuhl K, Kharasch ED, Ellingrod VL*, et al* (2013). Clinical Pharmacogenetics Implementation Consortium guideline for CYP2D6 and CYP2C19 genotypes and dosing of tricyclic antidepressants. *Clin Pharmacol Ther* **93**(5): 402-408.

Hiemke C, Baumann P, Bergemann N, Conca A, Dietmaier O, Egberts K*, et al* (2011). AGNP consensus guidelines for therapeutic drug monitoring in psychiatry: update 2011. *Pharmacopsychiatry* **44**(6): 195-235.

Hindorff L, MacArthur J, Morales J, Junkins H, Hall P, Klemm A*, et al* A Catalog of Published Genome-Wide Association Studies. . Vol 2014.

Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS*, et al* (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A* **106**(23): 9362-9367.

Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K (2002). A comprehensive review of genetic association studies. *Genet Med* **4**(2): 45-61.

Hodgson K, McGuffin P (2013). The genetic basis of depression. *Current topics in behavioral neurosciences* **14**: 81-99.

Hodgson K, Mufti SJ, Uher R, McGuffin P (2012). Genome-wide approaches to antidepressant treatment: working towards understanding and predicting response. *Genome Med* **4**(6): 52.

Horder J, Matthews P, Waldmann R (2011). Placebo, prozac and PLoS: significant lessons for psychopharmacology. *J Psychopharmacol* **25**(10): 1277-1288.

Horstmann S, Binder EB (2009a). Pharmacogenomics of antidepressant drugs. *Pharmacol Ther* **124**(1): 57-73.

Horstmann S, Dose T, Lucae S, Kloiber S, Menke A, Hennings J*, et al* (2009b). Suppressive effect of mirtazapine on the HPA system in acutely depressed women seems to be transient and not related to antidepressant action. *Psychoneuroendocrinology* **34**(2): 238-248.

Horvath S, Zhang B, Carlson M, Lu KV, Zhu S, Felciano RM*, et al* (2006). Analysis of oncogenic signaling networks in glioblastoma identifies ASPM as a molecular target. *Proceedings of the National Academy of Sciences* **103**(46): 17402-17407.

Howie B, Marchini J, Stephens M (2011). Genotype Imputation with Thousands of Genomes. *G3: Genes, Genomes, Genetics* **1**(6): 457-470.

Howie BN, Donnelly P, Marchini J (2009). A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. *PLoS Genet* **5**(6): e1000529.

Hu XZ, Lipsky RH, Zhu G, Akhtar LA, Taubman J, Greenberg BD*, et al* (2006). Serotonin transporter promoter gain-of-function genotypes are linked to obsessive-compulsive disorder. *Am J Hum Genet* **78**(5): 815-826.

Hu XZ, Rush AJ, Charney D, Wilson AF, Sorant AJ, Papanicolaou GJ*, et al* (2007). Association between a functional serotonin transporter promoter polymorphism and citalopram treatment in adult outpatients with major depression. *Arch Gen Psychiatry* **64**(7): 783-792.

Huezo-Diaz P, Perroud N, Spencer EP, Smith R, Sim S, Virding S*, et al* (2012). CYP2C19 genotype predicts steady state escitalopram concentration in GENDEP. *J Psychopharmacol* **26**(3): 398-407.

Huezo-Diaz P, Uher R, Smith R, Rietschel M, Henigsberg N, Marusic A*, et al* (2009). Moderation of antidepressant response by the serotonin transporter gene. *Br J Psychiatry* **195**(1): 30-38.

Iga J, Ueno S, Yamauchi K, Numata S, Kinouchi S, Tayoshi-Shibuya S*, et al* (2007a). Altered HDAC5 and CREB mRNA expressions in the peripheral leukocytes of major depression. *Progress in neuro-psychopharmacology & biological psychiatry* **31**(3): 628-632.

Iga J, Ueno S, Yamauchi K, Numata S, Tayoshi-Shibuya S, Kinouchi S*, et al* (2007b). Gene expression and association analysis of vascular endothelial growth factor in major depressive disorder. *Prog Neuropsychopharmacol Biol Psychiatry* **31**(3): 658-663.

Ingelman-Sundberg M (2004). Pharmacogenetics of cytochrome P450 and its applications in drug therapy: the past, present and future. *Trends Pharmacol Sci* **25**(4): 193-200.

Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG (2001). Replication validity of genetic association studies. *Nat Genet* **29**(3): 306-309.

Ioannidis JPA (2005). Why Most Published Research Findings Are False. *PLoS Med* **2**(8): e124.

Jablensky AV, Morgan V, Zubrick SR, Bower C, Yellachich LA (2005). Pregnancy, delivery, and neonatal complications in a population cohort of women with schizophrenia and major affective disorders. *Am J Psychiatry* **162**(1): 79-91.

Jansen RC, Nap J-P (2001). Genetical genomics: the added value from segregation. *Trends in Genetics* **17**(7): 388-391.

Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PIW (2008). SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* **24**(24): 2938-2939.

Jungke P, Ostrow G, Li JL, Norton S, Nieber K, Kelber O*, et al* (2011). Profiling of hypothalamic and hippocampal gene expression in chronically stressed rats treated with St. John's wort extract (STW 3-VI) and fluoxetine. *Psychopharmacology (Berl)* **213**(4): 757-772.

Kálmán J, Palotás A, Juhász A, Rimanóczy Á, Hugyecz M, Kovács Z*, et al* (2005). Impact of Venlafaxine on Gene Expression Profile in Lymphocytes of the Elderly with Major Depression – Evolution of Antidepressants and the Role of the “Neuro-Immune” System. *Neurochemical Research* **30**(11): 1429-1438.

Kangelaris KN, Bent S, Nussbaum RL, Garcia DA, Tice JA (2009). Genetic testing before anticoagulation? A systematic review of pharmacogenetic dosing of warfarin. *Journal of general internal medicine* **24**(5): 656-664.

Kato M, Serretti A (2010). Review and meta-analysis of antidepressant pharmacogenetic findings in major depressive disorder. *Molecular psychiatry* **15**(5): 473-500.

Kendler KS, Karkowski LM, Prescott CA (1999). Causal relationship between stressful life events and the onset of major depression. *Am J Psychiatry* **156**(6): 837-841.

Kennedy S, McIntyre R, Fallu A, Lam R (2002). Pharmacotherapy to sustain the fully remitted state. *J Psychiatry Neurosci* **27**(4): 269-280.

Kennedy SH, Eisfeld BS, Dickens SE, Bacchiochi JR, Bagby RM (2000). Antidepressant-induced sexual dysfunction during treatment with moclobemide, paroxetine, sertraline, and venlafaxine. *J Clin Psychiatry* **61**(4): 276-281.

Kent JT (1982). Robust properties of likelihood ratio tests. *Biometrika* **69**(1): 19-27.

Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE (2005). LIfetime prevalence and age-of-onset distributions of dsm-iv disorders in the national comorbidity survey replication. *Archives of General Psychiatry* **62**(6): 593-602.

Khan A, Leventhal RM, Khan SR, Brown WA (2002). Severity of depression and response to antidepressants and placebo: an analysis of the Food and Drug Administration database. *J Clin Psychopharmacol* **22**(1): 40-45.

Kim Y-K, Na K-S, Shin K-H, Jung H-Y, Choi S-H, Kim J-B (2007). Cytokine imbalance in the pathophysiology of major depressive disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* **31**(5): 1044-1053.

Kirchheiner J, Brosen K, Dahl ML, Gram LF, Kasper S, Roots I*, et al* (2001). CYP2D6 and CYP2C19 genotype-based dose recommendations for antidepressants: a first step towards subpopulation-specific dosages. *Acta Psychiatr Scand* **104**(3): 173-192.

Kirchheiner J, Nickchen K, Bauer M, Wong ML, Licinio J, Roots I*, et al* (2004). Pharmacogenetics of antidepressants and antipsychotics: the contribution of allelic variations to the phenotype of drug response. *Molecular psychiatry* **9**(5): 442-473.

Kirchheiner J, Rodriguez-Antona C (2009). Cytochrome P450 2D6 genotyping: potential role in improving treatment outcomes in psychiatric disorders. *CNS drugs* **23**(3): 181-191.

Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT (2008). Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Med* **5**(2): e45.

Knight J, Barnes MR, Breen G, Weale ME (2011). Using Functional Annotation for the Empirical Determination of Bayes Factors for Genome-Wide Association Study Analysis. *PLoS ONE* **6**(4): e14808.

Kuzman MR, Medved V, Bozina N, Hotujac L, Sain I, Bilusic H (2008). The influence of 5-HT(2C) and MDR1 genetic polymorphisms on antipsychotic-induced weight gain in female schizophrenic patients. *Psychiatry Res* **160**(3): 308-315.

Lamba JK, Lin YS, Thummel K, Daly A, Watkins PB, Strom S*, et al* (2002). Common allelic variants of cytochrome P4503A4 and their prevalence in different populations. *Pharmacogenetics and Genomics* **12**(2): 121-132.

Lane P (2008). Handling drop-out in longitudinal clinical trials: a comparison of the LOCF and MMRM approaches. *Pharmaceutical statistics* **7**(2): 93-106.

Langfelder P, Horvath S (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**(1): 559.

Lappalainen T, Sammeth M, Friedlander MR, t Hoen PA, Monlong J, Rivas MA*, et al* (2013). Transcriptome and genome sequencing uncovers functional variation in humans. *Nature* **501**(7468): 506-511.

Lecklin A, Etu-Seppala P, Stark H, Tuomisto L (1998). Effects of intracerebroventricularly infused histamine and selective H1, H2 and H3 agonists on food and water intake and urine flow in Wistar rats. *Brain Res* **793**(1-2): 279-288.

Lee B-H, Kim Y-K (2010). The Roles of BDNF in the Pathophysiology of Major Depression and in Antidepressant Treatment. *Psychiatry Investig* **7**(4): 231-235.

Lee HK, Hsu AK, Sajdak J, Qin J, Pavlidis P (2004). Coexpression analysis of human genes across many microarray data sets. *Genome research* **14**(6): 1085-1094.

Lee SH, Ripke S, Neale BM, Faraone SV, Purcell SM, Perlis RH*, et al* (2013). Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet* **45**(9): 984-994.

Leek JT, Storey JD (2007). Capturing Heterogeneity in Gene Expression Studies by Surrogate Variable Analysis. *PLoS Genet* **3**(9): e161.

Lesch KP, Balling U, Gross J, Strauss K, Wolozin BL, Murphy DL*, et al* (1994). Organization of the human serotonin transporter gene. *Journal of Neural Transmission* **95**(2): 157-162.

Lesch KP, Bengel D, Heils A, Sabol SZ, Greenberg BD, Petri S*, et al* (1996). Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* **274**(5292): 1527-1531.

Lesch KP, Schmitt A (2002). Antidepressants and gene expression profiling: how to SNARE novel drug targets. *Pharmacogenomics J* **2**(6): 346-348.

Li J, Ji L (2005). Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity* **95**(3): 221-227.

Li L, Kabesch M, Bouzigon E, Demenais F, Farrall M, Moffatt MF*, et al* (2013). Using eQTL weights to improve power for genome-wide association studies: a genetic study of childhood asthma. *Frontiers in genetics* **4**: 103.

Liang L, Morar N, Dixon AL, Lathrop GM, Abecasis GR, Moffatt MF*, et al* (2013). A cross-platform analysis of 14,177 expression quantitative trait loci derived from lymphoblastoid cell lines. *Genome research* **23**(4): 716-726.

Liew CC, Ma J, Tang HC, Zheng R, Dempsey AA (2006). The peripheral blood transcriptome dynamically reflects system wide biology: a potential diagnostic tool. *J Lab Clin Med* **147**(3): 126-132.

Lingjaerde O, Ahlfors UG, Bech P, Dencker SJ, Elgen K (1987). The UKU side effect rating scale. A new comprehensive rating scale for psychotropic drugs and a cross-sectional study of side effects in neuroleptic-treated patients. *Acta Psychiatr Scand Suppl* **334**: 1-100.

Lourdusamy A, Newhouse S, Lunnon K, Proitsi P, Powell J, Hodges A*, et al* (2012). Identification of cis-regulatory variation influencing protein abundance levels in human plasma. *Hum Mol Genet*.

Lubke GH, Hottenga JJ, Walters R, Laurin C, de Geus EJC, Willemsen G*, et al* (2012). Estimating the Genetic Variance of Major Depressive Disorder Due to All Single Nucleotide Polymorphisms. *Biological Psychiatry* **72**(8): 707-709.

Lunnon K, Ibrahim Z, Proitsi P, Lourdusamy A, Newhouse S, Sattlecker M*, et al* (2012). Mitochondrial dysfunction and immune activation are detectable in early Alzheimer's disease blood. *J Alzheimers Dis* **30**(3): 685-710.

Mäkinen V-P, Civelek M, Meng Q, Zhang B, Zhu J, Levian C*, et al* (2014). Integrative Genomics Reveals Novel Molecular Pathways and Gene Networks for Coronary Artery Disease. *PLoS Genet* **10**(7): e1004502.

Mallinckrodt CH, Clark WS, David SR (2001). Accounting for dropout bias using mixed-effects models. *J Biopharm Stat* **11**(1-2): 9-21.

Mamdani F, Berlim MT, Beaulieu M-M, Turecki G (2013). Pharmacogenomic predictors of citalopram treatment outcome in major depressive disorder. *The World Journal of Biological Psychiatry* **15**(2): 135-144.

Mamdani F, Berlim MT, Beaulieu MM, Labbe A, Merette C, Turecki G (2011). Gene expression biomarkers of response to citalopram treatment in major depressive disorder. *Transl Psychiatry* **1**: e13.

Manji HK, Quiroz JA, Sporn J, Payne JL, Denicoff K, N AG*, et al* (2003). Enhancing neuronal plasticity and cellular resilience to develop novel, improved therapeutics for difficult-to-treat depression. *Biol Psychiatry* **53**(8): 707-742.

March JS, Silva SG, Compton S, Shapiro M, Califf R, Krishnan R (2005). The case for practical clinical trials in psychiatry. *Am J Psychiatry* **162**(5): 836-846.

Masand PS (2003). Tolerability and adherence issues in antidepressant therapy. *Clinical Therapeutics* **25**(8): 2289-2304.

Matchar DB, Thakur M (2007a). Is genetic testing for cytochrome P450 polymorphisms ready for implementation? *American family physician* **76**(3): 348, 351.

Matchar DB, Thakur ME, Grossman I, McCrory DC, Orlando LA, Steffens DC*, et al* (2007b). Testing for cytochrome P450 polymorphisms in adults with non-psychotic depression treated with selective serotonin reuptake inhibitors (SSRIs). *Evid Rep Technol Assess (Full Rep)*, 2007/09/04 edn, pp 1-77.

McGuffin P, Cohen S, Knight J (2007). Homing in on depression genes. *Am J Psychiatry* **164**(2): 195-197.

McKenzie M, Henders AK, Caracella A, Wray NR, Powell JE (2014). Overlap of expression quantitative trait loci (eQTL) in human brain and blood. *BMC Med Genomics* **7**: 31.

Meijer WE, Heerdink ER, van Eijk JT, Leufkens HG (2002). Adverse events in users of sertraline: results from an observational study in psychiatric practice in The Netherlands. *Pharmacoepidemiol Drug Saf* **11**(8): 655-662.

Miller JA, Oldham MC, Geschwind DH (2008). A Systems Level Analysis of Transcriptional Changes in Alzheimer's Disease and Normal Aging. *The Journal of Neuroscience* **28**(6): 1410-1420.

Mitchell AJ (2006). Depressed patients and treatment adherence. *Lancet* **367**(9528): 2041-2043.

Montgomery SA, Asberg M (1979). A new depression scale designed to be sensitive to change. *Br J Psychiatry* **134**: 382-389.

Monti JM, Orellana C, Boussard M, Jantos H, Olivera S (1990). Sleep variables are unaltered by zolantidine in rats: are histamine H2-receptors not involved in sleep regulation? *Brain Res Bull* **25**(2): 229-231.

Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, Spielman RS*, et al* (2004). Genetic analysis of genome-wide variation in human gene expression. *Nature* **430**(7001): 743-747.

Morris AP, Voight BF, Teslovich TM, Ferreira T, Segre AV, Steinthorsdottir V*, et al* (2012). Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet* **44**(9): 981-990.

Moskvina V, Farmer A, Swainson V, O'Leary J, Gunasinghe C, Owen M*, et al* (2007). Interrelationship of childhood trauma, neuroticism, and depressive phenotype. *Depress Anxiety* **24**(3): 163-168.

Mrazek DA, Biernacka JM, O'Kane DJ, Black JL, Cunningham JM, Drews MS*, et al* (2011). CYP2C19 variation and citalopram response. *Pharmacogenetics and genomics* **21**(1): 1-9.

Murphy GM, Jr., Hollander SB, Rodrigues HE, Kremer C, Schatzberg AF (2004). Effects of the serotonin transporter gene promoter polymorphism on mirtazapine and paroxetine efficacy and adverse events in geriatric major depression. *Arch Gen Psychiatry* **61**(11): 1163-1169.

Murphy GM, Jr., Pollock BG, Kirshner MA, Pascoe N, Cheuk W, Mulsant BH*, et al* (2001). CYP2D6 genotyping with oligonucleotide microarrays and nortriptyline concentrations in geriatric depression. *Neuropsychopharmacology* **25**(5): 737-743.

Nanni V, Uher R, Danese A (2012). Childhood maltreatment predicts unfavorable course of illness and treatment outcome in depression: a meta-analysis. *The American journal of psychiatry* **169**(2): 141-151.

Nelson JC (1999). A review of the efficacy of serotonergic and noradrenergic reuptake inhibitors for treatment of major depression. *Biological Psychiatry* **46**(9): 1301-1308.

Ng C, Sarris J, Singh A, Bousman C, Byron K, Peh LH*, et al* (2013). Pharmacogenetic polymorphisms and response to escitalopram and venlafaxine over 8 weeks in major depression. *Human Psychopharmacology: Clinical and Experimental* **28**(5): 516-522.

Nica AC, Parts L, Glass D, Nisbet J, Barrett A, Sekowska M*, et al* (2011). The Architecture of Gene Regulatory Variation across Multiple Human Tissues: The MuTHER Study. *PLoS Genet* **7**(2): e1002003.

Nicholson G, Rantalainen M, Li JV, Maher AD, Malmodin D, Ahmadi KR*, et al* (2011). A Genome-Wide Metabolic QTL Analysis in Europeans Implicates Two Loci Shaped by Recent Positive Selection. *PLoS Genet* **7**(9): e1002270.

Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ (2010). Trait-Associated SNPs Are More Likely to Be eQTLs: Annotation to Enhance Discovery from GWAS. *PLoS Genet* **6**(4): e1000888.

Nikisch G, Mathe AA, Czernik A, Eap CB, Jimenez-Vasquez P, Brawand-Amey M*, et al* (2004). Stereoselective metabolism of citalopram in plasma and cerebrospinal fluid of depressive patients: relationship with 5-HIAA in CSF and clinical response. *J Clin Psychopharmacol* **24**(3): 283-290.

Nordin C, Bertilsson L (1995). Active Hydroxymetabolites of Antidepressants. *Clin Pharmacokinet* **28**(1): 26-40.

Nutt DJ (2002). Tolerability and safety aspects of mirtazapine. *Human Psychopharmacology: Clinical and Experimental* **17**(S1): S37-S41.

Nutt DJ, Sharpe M (2008). Uncritical positive regard? Issues in the efficacy and safety of psychotherapy. *J Psychopharmacol* **22**(1): 3-6.

Nyholt DR (2004). A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet* **74**(4): 765-769.

O'Reilly RL, Bogue L, Singh SM (1994). Pharmacogenetic response to antidepressants in a multicase family with affective disorder. *Biol Psychiatry* **36**(7): 467-471.

Obeid LM, Linardic CM, Karolak LA, Hannun YA (1993). Programmed cell death induced by ceramide. *Science* **259**(5102): 1769-1771.

Oldham MC, Konopka G, Iwamoto K, Langfelder P, Kato T, Horvath S*, et al* (2008). Functional organization of the transcriptome in human brain. *Nat Neurosci* **11**(11): 1271-1282.

Olesen OV, Linnet K (1997a). Hydroxylation and Demethylation of the Tricyclic Antidepressant Nortriptyline by cDNA-Expressed Human Cytochrome P-450 Isozymes. *Drug Metabolism and Disposition* **25**(6): 740-744.

Olesen OV, Linnet K (1997b). Metabolism of the tricyclic antidepressant amitriptyline by cDNA-expressed human cytochrome P450 enzymes. *Pharmacology* **55**(5): 235-243.

Olesen OV, Linnet K (1999). Studies on the stereoselective metabolism of citalopram by human liver microsomes and cDNA-expressed cytochrome P450 enzymes. *Pharmacology* **59**(6): 298-309.

Ostad Haji E, Tadic A, Wagner S, Dragicevic A, Müller MJ, Boland K*, et al* (2011). Association Between Citalopram Serum Levels and Clinical Improvement of Patients With Major Depression. *Journal of Clinical Psychopharmacology* **31**(3): 281-286 210.1097/JCP.1090b1013e318218f318503.

Palotás M, oacute, Palotás A, aacute, Puskás L, aacute*, et al* (2004). Gene expression profile analysis of the rat cortex following treatment with imipramine and citalopram. *The International Journal of Neuropsychopharmacology* **7**(04): 401-413.

Pariante CM, Lightman SL (2008). The HPA axis in major depression: classical theories and new developments. *Trends in neurosciences* **31**(9): 464-468.

Pe'er I, Yelensky R, Altshuler D, Daly MJ (2008). Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genetic epidemiology* **32**(4): 381-385.

Perlis RH, Laje G, Smoller JW, Fava M, Rush AJ, McMahon FJ (2009). Genetic and clinical predictors of sexual dysfunction in citalopram-treated depressed patients. *Neuropsychopharmacology* **34**(7): 1819-1828.

Perry PJ, Zeilmann C, Arndt S (1994). Tricyclic antidepressant concentrations in plasma: an estimate of their sensitivity and specificity as a predictor of response. *J Clin Psychopharmacol* **14**(4): 230-240.

Peters EJ, Slager SL, Kraft JB, Jenkins GD, Reinalda MS, McGrath PJ*, et al* (2008). Pharmacokinetic genes do not influence response or tolerance to citalopram in the STAR D sample. *PloS one* **3**(4): e1872.

Petretto E, Mangion J, Dickens NJ, Cook SA, Kumaran MK, Lu H*, et al* (2006). Heritability and tissue specificity of expression quantitative trait loci. *PLoS Genet* **2**(10): e172.

Pickrell Joseph K (2014). Joint Analysis of Functional Genomic Data and Genome-wide Association Studies of 18 Human Traits. *The American Journal of Human Genetics* **94**(4): 559-573.

Pirooznia M, Yang JY, Yang MQ, Deng Y (2008). A comparative study of different machine learning methods on microarray gene expression data. *BMC Genomics* **9 Suppl 1**: S13.

Pittenger C, Duman RS (2008). Stress, depression, and neuroplasticity: a convergence of mechanisms. *Neuropsychopharmacology* **33**(1): 88-109.

Plagnol V, Smyth DJ, Todd JA, Clayton DG (2009). Statistical independence of the colocalized association signals for type 1 diabetes and RPS26 gene expression on chromosome 12q13. *Biostatistics (Oxford, England)* **10**(2): 327-334.

Powell TR, Tansey KE, Breen G, Farmer AE, Craig IW, Uher R*, et al* (2013). ATP-binding cassette sub-family F member 1 (ABCF1) is identified as a putative therapeutic target of escitalopram in the inflammatory cytokine pathway. *J Psychopharmacol* **27**(7): 609-615.

Preskorn SH, Fast GA (1991). Therapeutic drug monitoring for antidepressants: efficacy, safety, and cost effectiveness. *J Clin Psychiatry* **52 Suppl**: 23-33.

Price AL, Helgason A, Thorleifsson G, McCarroll SA, Kong A, Stefansson K (2011). Single-tissue and cross-tissue heritability of gene expression via identity-by-descent in related or unrelated individuals. *PLoS Genet* **7**(2): e1001317.

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38**(8): 904-909.

Prusoff BA, Klerman GL, Paykel ES (1972). COncordance between clinical assessments and patients' self-report in depression. *Archives of General Psychiatry* **26**(6): 546-552.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D*, et al* (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**(3): 559-575.

Quinlan AR, Hall IM (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**(6): 841-842.

Ramasamy A, Trabzuni D, Gibbs JR, Dillman A, Hernandez DG, Arepalli S*, et al* (2013). Resolving the polymorphism-in-probe problem is critical for correct interpretation of expression QTL studies. *Nucleic acids research* **41**(7): e88.

Rasmussen BB, Brosen K (2000). Is therapeutic drug monitoring a case for optimizing clinical outcome and avoiding interactions of the selective serotonin reuptake inhibitors? *Ther Drug Monit* **22**(2): 143-154.

Rebsamen MC, Desmeules J, Daali Y, Chiappe A, Diemand A, Rey C*, et al* (2009). The AmpliChip CYP450 test: cytochrome P450 2D6 genotype assessment and phenotype prediction. *Pharmacogenomics J* **9**(1): 34-41.

Reis M, Aamo T, Spigset O, Ahlner J (2009). Serum concentrations of antidepressant drugs in a naturalistic setting: compilation based on a large therapeutic drug monitoring database. *Ther Drug Monit* **31**(1): 42-56.

Reis M, Cherma MD, Carlsson B, Bengtsson F (2007). Therapeutic drug monitoring of escitalopram in an outpatient setting. *Ther Drug Monit* **29**(6): 758-766.

Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM, Breen G*, et al* (2013). A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry* **18**(4): 497-511.

Risselada AJ, Vehof J, Bruggeman R, Wilffert B, Cohen D, Al Hadithy AF*, et al* (2010). Association between HTR2C gene polymorphisms and the metabolic syndrome in patients using antipsychotics: a replication study. *Pharmacogenomics J*.

Rudberg I, Hendset M, Uthus LH, Molden E, Refsum H (2006). Heterozygous Mutation in CYP2C19 Significantly Increases the Concentration/Dose Ratio of Racemic Citalopram and Escitalopram (S-citalopram). *Ther Drug Monit* **28**(1): 102-105.

Rudberg I, Mohebi B, Hermann M, Refsum H, Molden E (2008). Impact of the ultrarapid CYP2C19*17 allele on serum concentration of escitalopram in psychiatric patients. *Clin Pharmacol Ther* **83**(2): 322-327.

Rush AJ (2007). The varied clinical presentations of major depressive disorder. *J Clin Psychiatry* **68 Suppl 8**: 4-10.

Rush AJ, Fava M, Wisniewski SR, Lavori PW, Trivedi MH, Sackeim HA*, et al* (2004). Sequenced treatment alternatives to relieve depression (STAR*D): rationale and design. *Controlled clinical trials* **25**(1): 119-142.

San L, Arranz B (2008). Agomelatine: a novel mechanism of antidepressant action involving the melatonergic and the serotonergic system. *Eur Psychiatry* **23**(6): 396-402.

Sanchez C, Bergqvist PB, Brennum LT, Gupta S, Hogg S, Larsen A*, et al* (2003). Escitalopram, the S-(+)-enantiomer of citalopram, is a selective serotonin reuptake inhibitor with potent effects in animal models predictive of antidepressant and anxiolytic activities. *Psychopharmacology (Berl)* **167**(4): 353-362.

Sanchez C, Hyttel J (1999). Comparison of the effects of antidepressants and their metabolites on reuptake of biogenic amines and on receptor binding. *Cell Mol Neurobiol* **19**(4): 467-489.

Schadt EE, Monks SA, Drake TA, Lusis AJ, Che N, Colinayo V*, et al* (2003). Genetics of gene expression surveyed in maize, mouse and man. *Nature* **422**(6929): 297-302.

Schizophrenia Working Group of the Psychiatric Genomics C (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**(7510): 421-427.

Seldin MF, Shigeta R, Villoslada P, Selmi C, Tuomilehto J, Silva G*, et al* (2006). European population substructure: clustering of northern and southern populations. *PLoS Genet* **2**(9): e143.

Serretti A, Calati R, Massat I, Linotte S, Kasper S, Lecrubier Y*, et al* (2009). Cytochrome P450 CYP1A2, CYP2C9, CYP2C19 and CYP2D6 genes are not associated with response and remission in a sample of depressive patients. *International Clinical Psychopharmacology* **24**(5): 250-256 210.1097/YIC.1090b1013e32832e32835b32830d.

Shabalin AA (2012). Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics* **28**(10): 1353-1358.

Sharp T, Cowen PJ (2011). 5-HT and depression: is the glass half-full? *Current opinion in pharmacology* **11**(1): 45-51.

Sillaber I, Panhuysen M, Henniger MS, Ohl F, Kuhne C, Putz B*, et al* (2008). Profiling of behavioral changes and hippocampal gene expression in mice chronically treated with the SSRI paroxetine. *Psychopharmacology (Berl)* **200**(4): 557-572.

Sim SC, Risinger C, Dahl ML, Aklillu E, Christensen M, Bertilsson L*, et al* (2006). A common novel CYP2C19 gene variant causes ultrarapid drug metabolism relevant for the drug response to proton pump inhibitors and antidepressants. *Clinical pharmacology and therapeutics* **79**(1): 103-113.

Small KS, Hedman AK, Grundberg E, Nica AC, Thorleifsson G, Kong A*, et al* (2011). Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. *Nat Genet* **43**(6): 561-564.

Smemo S, Tena JJ, Kim K-H, Gamazon ER, Sakabe NJ, Gomez-Marin C*, et al* (2014). Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature* **advance online publication**.

Smyth GK (2005). limma: Linear Models for Microarray Data. In: Gentleman R, Carey V, Huber W, Irizarry R, Dudoit S, Smyth GK (eds). *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Springer New York, pp 397-420.

Solovyev VV, Shahmuradov IA (2003). PromH: Promoters identification using orthologous genomic sequences. *Nucleic acids research* **31**(13): 3540-3545.

Stahl SM (1998). Selecting an antidepressant by using mechanism of action to enhance efficacy and avoid side effects. *J Clin Psychiatry* **59 Suppl 18**: 23-29.

StataCorp. (2009). Stata Statistical Software: Release 11. . StataCorp LP.: College Station, TX.

Steimer W, Zopf K, von Amelunxen S, Pfeiffer H, Bachofer J, Popp J*, et al* (2005). Amitriptyline or not, that is the question: pharmacogenetic testing of CYP2D6 and CYP2C19 identifies patients with low or high risk for side effects in amitriptyline therapy. *Clin Chem* **51**(2): 376-385.

Stranger BE, Nica AC, Forrest MS, Dimas A, Bird CP, Beazley C*, et al* (2007). Population genomics of human gene expression. *Nat Genet* **39**(10): 1217-1224.

Streiner DL (2002). Breaking up is hard to do: the heartbreak of dichotomizing continuous data. *Canadian journal of psychiatry Revue canadienne de psychiatrie* **47**(3): 262-266.

Strohmaier J, Wust S, Uher R, Henigsberg N, Mors O, Hauser J*, et al* (2011). Sexual dysfunction during treatment with serotonergic and noradrenergic antidepressants: Clinical description and the role of the 5-HTTLPR. *World J Biol Psychiatry*.

Sugai T, Suzuki Y, Sawamura K, Fukui N, Inoue Y, Someya T (2006). The effect of 5-hydroxytryptamine 3A and 3B receptor genes on nausea induced by paroxetine. *Pharmacogenomics J* **6**(5): 351-356.

Sullivan PF, Fan C, Perou CM (2006). Evaluating the comparability of gene expression in blood and brain. *Am J Med Genet B Neuropsychiatr Genet* **141B**(3): 261-268.

Sullivan PF, Neale MC, Kendler KS (2000). Genetic epidemiology of major depression: review and meta-analysis. *Am J Psychiatry* **157**(10): 1552-1562.

Tanaka M, Kobayashi D, Murakami Y, Ozaki N, Suzuki T, Iwata N*, et al* (2008). Genetic polymorphisms in the 5-hydroxytryptamine type 3B receptor gene and paroxetine-induced nausea. *Int J Neuropsychopharmacol* **11**(2): 261-267.

Tansey K, Hodgson K, Powell TR, Coppola G, Hauser J, Henigsberg N*, et al* (in prep). Transcriptomic signatures predict response to antidepressant treatment in major depressive disorder.

Tansey KE, Guipponi M, Hu X, Domenici E, Lewis G, Malafosse A*, et al* (2013). Contribution of common genetic variants to antidepressant response. *Biol Psychiatry* **73**(7): 679-682.

Tansey KE, Guipponi M, Perroud N, Bondolfi G, Domenici E, Evans D*, et al* (2012). Genetic Predictors of Response to Serotonergic and Noradrenergic Antidepressants in Major Depressive Disorder: A Genome-Wide Analysis of Individual-Level Data and a Meta-Analysis. *PLoS Med* **9**(10): e1001326.

Theisen FM, Hinney A, Bromel T, Heinzel-Gutenbrunner M, Martin M, Krieg JC*, et al* (2004). Lack of association between the -759C/T polymorphism of the 5-HT2C receptor gene and clozapine-induced weight gain among German schizophrenic individuals. *Psychiatr Genet* **14**(3): 139-142.

Thomas L, Mulligan J, Mason V, Tallon D, Wiles N, Cowen P*, et al* (2008). GENetic and clinical predictors of treatment response in depression: the GenPod randomised trial protocol. *Trials* **9**: 29.

Tian C, Plenge RM, Ransom M, Lee A, Villoslada P, Selmi C*, et al* (2008). Analysis and application of European genetic substructure using 300 K SNP information. *PLoS Genet* **4**(1): e4.

Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, Warden D, Ritz L*, et al* (2006). Evaluation of outcomes with citalopram for depression using measurement-based care in STAR D: implications for clinical practice. *The American journal of psychiatry* **163**(1): 28-40.

Tsai MH, Lin KM, Hsiao MC, Shen WW, Lu ML, Tang HS*, et al* (2010). Genetic polymorphisms of cytochrome P450 enzymes influence metabolism of the antidepressant escitalopram and treatment response. *Pharmacogenomics* **11**(4): 537-546.

Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R (2008). Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy. *New England Journal of Medicine* **358**(3): 252-260.

Uher R, Farmer A, Henigsberg N, Rietschel M, Mors O, Maier W*, et al* (2009a). Adverse reactions to antidepressants. *The British journal of psychiatry : the journal of mental science* **195**(3): 202-210.

Uher R, Farmer A, Maier W, Rietschel M, Hauser J, Marusic A*, et al* (2008). Measuring depression: comparison and integration of three scales in the GENDEP study. *Psychol Med* **38**(2): 289-300.

Uher R, Huezo-Diaz P, Perroud N, Smith R, Rietschel M, Mors O*, et al* (2009b). Genetic predictors of response to antidepressants in the GENDEP project. *Pharmacogenomics J* **9**(4): 225-233.

Uher R, Maier W, Hauser J, MaruÅ¡iÄ A, Schmael C, Mors O*, et al* (2009c). Differential efficacy of escitalopram and nortriptyline on dimensional measures of depression. *The British Journal of Psychiatry* **194**(3): 252-259.

Uher R, Mors O, Rietschel M, Rajewska-Rager A, Petrovic A, Zobel A*, et al* (2011). Early and delayed onset of response to antidepressants in individual trajectories of change during treatment of major depression: a secondary analysis of data from the Genome-Based Therapeutic Drugs for Depression (GENDEP) study. *J Clin Psychiatry* **72**(11): 1478-1484.

Uher R, Perroud N, Ng MY, Hauser J, Henigsberg N, Maier W*, et al* (2010). Genome-wide pharmacogenetics of antidepressant response in the GENDEP project. *Am J Psychiatry* **167**(5): 555-564.

Uher R, Tansey KE, Malki K, Perlis RH (2012). Biomarkers predicting treatment outcome in depression: what is clinically significant? *Pharmacogenomics* **13**(2): 233-240.

Uhr M, Steckler T, Yassouridis A, Holsboer F (2000). Penetration of Amitriptyline, but Not of Fluoxetine, into Brain is Enhanced in Mice with Blood-Brain Barrier Deficiency Due to Mdr1a P-Glycoprotein Gene Disruption. *Neuropsychopharmacology* **22**(4): 380-387.

van Nas A, Ingram-Drake L, Sinsheimer JS, Wang SS, Schadt EE, Drake T*, et al* (2010). Expression Quantitative Trait Loci: Replication, Tissue- and Sex-Specificity in Mice. *Genetics* **185**(3): 1059-1068.

Vandel S, Vandel B, Sandoz M, Allers G, Bechtel P, Volmat R (1978). Clinical response and plasma concentration of amitriptyline and its metabolite nortriptyline. *Eur J Clin Pharmacol* **14**(3): 185-190.

Veyrieras J-B, Kudaravalli S, Kim SY, Dermitzakis ET, Gilad Y, Stephens M*, et al* (2008). High-Resolution Mapping of Expression-QTLs Yields Insight into Human Gene Regulation. *PLoS Genet* **4**(10): e1000214.

von Moltke LL, Greenblatt DJ, Giancarlo GM, Granda BW, Harmatz JS, Shader RI (2001). Escitalopram (S-Citalopram) and Its Metabolites in Vitro: Cytochromes Mediating Biotransformation, Inhibitory Effects, and Comparison to R-Citalopram. *Drug Metabolism and Disposition* **29**(8): 1102-1109.

Ward LD, Kellis M (2012). HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic acids research* **40**(Database issue): D930-934.

Westra HJ, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J*, et al* (2013). Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* **45**(10): 1238-1243.

Whyte IM, Dawson AH, Buckley NA (2003). Relative toxicity of venlafaxine and selective serotonin reuptake inhibitors in overdose compared to tricyclic antidepressants. *QJM* **96**(5): 369-374.

Wing JK, Sartorius N, Üstün TB (1998). *Diagnosis and clinical measurement in psychiatry: a reference manual for SCAN* Cambridge University Press.

Wittke-Thompson JK, Pluzhnikov A, Cox NJ (2005). Rational inferences about departures from Hardy-Weinberg equilibrium. *Am J Hum Genet* **76**(6): 967-986.

Wong EHF, Sonders MS, Amara SG, Tinholt PM, Piercey MFP, Hoffmann WP*, et al* (2000). Reboxetine: a pharmacologically potent, selective, and specific norepinephrine reuptake inhibitor. *Biological Psychiatry* **47**(9): 818-829.

World Health Organisation (1992). *The ICD-10 Classification of Mental Health and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines* WHO.

World Health Organisation (2008). The global burden of disease: 2004 update. World Health Organisation.

Xiao R, Boehnke M (2009). Quantifying and correcting for the winner's curse in genetic association studies. *Genetic epidemiology* **33**(5): 453-462.

Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR*, et al* (2010). Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* **42**(7): 565-569.

Yao C, Joehanes R, Johnson AD, Huan T, Esko T, Ying S*, et al* (2014). Sex- and age-interacting eQTLs in human complex diseases. *Hum Mol Genet* **23**(7): 1947-1956.

Zimmerman M, Posternak MA, Chelminski I (2004). Implications of using different cut-offs on symptom severity scales to define remission from depression. *International Clinical Psychopharmacology* **19**(4): 215-220.

Chapter 10 Appendices

Appendix A. Montgomery Asberg Depression Rating Scale

GENDEP

GENDEP Unique ID		Date Completed:	
Researcher Name:		Week Number:	
Rating Score:		Location:	Clinic /Centre Home Telephone

Montgomery and Asberg Depression Rating Scale

The rating should be based on a clinical interview moving from broadly phrased questions about symptoms to more detailed ones which allow a precise rating of severity. The rater must decide whether the rating lies on the defined scale steps (0, 2, 4, 6) or between them (1, 3, 5) and then report the appropriate number. It is important that it is only on rare occasions that a depressed patient is encountered who cannot be rated on items in the scale. If definite items can not be elicited from the patient all relevant clues as well as information from other sources should be used as a basis for the rating in line with customary clinical practice. The items should be rated with regards to the state of the patient over the past week.

---

1 - APPARENT SADNESS - Representing despondency, gloom and despair, (more than just ordinary transient low spirits) reflected in speech, facial expression, and posture. Rate by depth and inability to brighten up.

0 No sadness
1
2 Looks dispirited but does brighten up without difficulty
3
4 Appears sad and unhappy most of the time
5
6 Looks miserable all the time. Extremely despondent.

---

2 - REPORTED SADNESS - *Representing reports of depressed mood, regardless of whether it is reflected in appearance or not. Includes low spirits, despondency or the feeling of being beyond help and without hope. Rate according to intensity, duration and the extent to which the mood is reported to be influenced by events.*

0 Occasional sadness in keeping with the circumstances.
1
2 Sad or low but brightens up without difficulty.
3
4 Pervasive feelings of sadness or gloominess. The mood is still influenced by external circumstances.
5
6 Continuous or unvarying sadness, misery or despondency.

---

3 - INNER TENSION - *Representing feelings of ill-defined discomfort, edginess, inner turmoil, mental tension mounting to either panic, dread or anguish. Rate according to intensity, frequency, duration and the extent of reassurance called for.*

0 Placid. Only fleeting inner tension.

1

2 Occasional feelings of edginess and ill-defined discomfort

3

4 Continuous feelings of inner tension or intermittent panic which the patient can only master with some difficulty.

5

6 Unrelenting dread or anguish. Overwhelming panic.

---

4 - REDUCED SLEEP - *Representing the experience of reduced duration or depth of sleep compared to the subject's own normal pattern when well.*

0 Sleeps as usual.

1

2 Slight difficulty dropping off to sleep or slightly reduced, light or fitful sleep

3

4 Sleep reduced or broken by at least two hours.

5

6 Less than two or three hours sleep.

---

5 - REDUCED APPETITE - *Representing the feeling of a loss of appetite compared with when well. Rate by loss of desire for food or the need to force oneself to eat.*

0 Normal or increased appetite.

1

2 Slightly reduced appetite

3

4 No appetite. Food is tasteless.

5

6 Needs persuasion to eat at all.

---

6 - CONCENTRATION DIFFICULTIES - *Representing difficulties in collecting one's thoughts mounting to incapacitating lack of concentration. Rate according to intensity, frequency, and degree of incapacity produced.*

0 No difficulties in concentrating.

1

2 Occasional difficulties in collecting one's thoughts.

3

4 Difficulties in concentrating and sustaining thought which reduces ability to read or hold a conversation.

5

6 Unable to read or converse without great difficulty.

---

7 - LASSITUDE - *Representing a difficulty getting started or slowness initiating and performing everyday activities.*

0 Hardly any difficulties in getting started. No sluggishness.

1

2 Difficulties in starting activities.

3

4 Difficulties in starting simple routine activities, which are carried out with effort.

5

6 Complete lassitude. Unable to do anything without help.

8 - INABILITY TO FEEL - *Representing the subjective experience of reduced interest in the surroundings, or activities that normally give pleasure.The ability to react with adequate emotion to circumstances or people is reduced.*

0 Normal interest in the surroundings and in other people.

1

2 Reduced ability to enjoy usual interests.

3

4 Loss of interest in the surroundings. Loss of feelings for friends and acquaintances.

5

6 The experience of being emotionally paralyzed, inability to feel anger, grief or pleasure and a complete or even painful failure to feel for close relatives and friends.

9 - PESSIMISTIC THOUGHTS - *Representing thoughts of guilt, inferiority, self-reproach, sinfulness, remorse and ruin.*

0 No pessimistic thoughts.

1

2 Fluctuating ideas of failure, self-reproach or self-depreciation.

3

4 Persistent self-accusations, or definite but still rational ideas of guilt or sin. Increasingly pessimistic about the future.

5

6 Delusions of ruin, remorse and unredeemable sin. Self-accusations which are absurd and unshakable.

10 - SUICIDAL THOUGHTS - *Representing the feeling that life is not worth living, that a natural death would be welcome, suicidal thoughts, and preparations for suicide. Suicidal attempts should not in themselves influence the rating.*

0 Enjoys life or takes it as it comes.

1

2 Weary of life. Only fleeting suicidal thoughts.

3

4 Probably better off dead. Suicidal thoughts are common, and suicide is considered as a possible solution, but without specific plans or intention.

5

6 Explicit plans for suicide when there is an opportunity. Active preparations for suicide.

Appendix B. Antidepressant Side Effects Checklist.

GENDEP

GENDEP Unique ID		Date Completed:	
Researcher Name:		Week Number:	
Medication (either A/B or other drug):		Location:	Clinic /Centre Home Telephone

Antidepressant Side-Effects Checklist

Score the following list of symptoms (0-3; 0 = absent, 1 = mild, 2 = moderate, 3 = severe)

Please indicate if the symptom is *likely to be* a side effect of medication (either study drug, or another antidepressant). Write a comment to provide relevant information if the item is **not** a side effect.

	List of symtoms	Score (0-3)				Linked to study drug?		Comment
1	Dry mouth	0	1	2	3	Y	N	
2	Drowsiness	0	1	2	3	Y	N	
3	Insomnia (difficulty sleeping)	0	1	2	3	Y	N	
4	Blurred vision	0	1	2	3	Y	N	
5	Headache	0	1	2	3	Y	N	
6	Constipation	0	1	2	3	Y	N	
7	Diarrhoea	0	1	2	3	Y	N	
8	Increased appetite	0	1	2	3	Y	N	
9	Decreased appetite	0	1	2	3	Y	N	
10	Nausea or Vomiting	0	1	2	3	Y	N	
	1 = slight nausea, 2 = more nausea but no vomiting, 3 = with vomiting							
11	Problems with urination	0	1	2	3	Y	N	
12	Problems with sexual function	0	1	2	3	Y	N	
13	Palpitations	0	1	2	3	Y	N	
14	Feeling light-headed on standing	0	1	2	3	Y	N	
15	Feeling like the room is spinning around	0	1	2	3	Y	N	
16	Sweating	0	1	2	3	Y	N	
17	Increased body temperature	0	1	2	3	Y	N	
18	Tremor	0	1	2	3	Y	N	
19	Disorientation	0	1	2	3	Y	N	
20	Yawning	0	1	2	3	Y	N	
21	Weight gain	0	1	2	3	Y	N	
						Score:		

Appendix C. Modified Toronto Side Effects Scale

Your symptoms

The following questions ask you about symptoms you may have experienced **during the past week**. Please tick only **one** answer for each question.

On how many days have you had……	None	1-3 days	4-7 days
1. Back pain	1	2	3
2. Chest pain	1	2	3
3. Stiffness in your arms or legs	1	2	3
4. Headaches	1	2	3
5. Sore throat	1	2	3
6. Tenderness of the glands in your neck	1	2	3
7. A rapid heart beat	1	2	3
8. Tremor	1	2	3
9. Agitation	1	2	3
10. Dry mouth	1	2	3
11. Excessive sweating	1	2	3
12. Tingling in your limbs, fingers or toes	1	2	3
13. Stomach pains	1	2	3
14. Constipation	1	2	3
15. Diarrhoea	1	2	3
16. Felt sick or nauseous	1	2	3
17. Noticed changes in the way your food tastes	1	2	3
18. Daytime drowsiness	1	2	3
19. Light headedness or dizziness	1	2	3
20. Shortness of breath at rest	1	2	3

On how many days have you had……	None	1-3 days	4-7 days
21. Ringing in the ears	1	2	3
22. Increased sensitivity to light or noise	1	2	3
23. Difficulty or pain passing urine	1	2	3
24. Passed urine more often	1	2	3
25. Swelling of your breasts	1	2	3
26. Skin rash or irritation	1	2	3
27. Hot flushes	1	2	3
28. Difficulty sleeping	1	2	3
If you have had any **OTHER PHYSICAL SYMPTOMS** during the past week that are not included in the above list, please specify below			
29.	1	2	3
30.	1	2	3

FOR MEN ONLY

In the past week, have you had any……	Yes	No	Not applicable
31. Difficulty ejaculating	1	2	3
32. Impotence (difficulty getting/maintaining an erection)	1	2	3

FOR WOMEN ONLY

In the past month, have you had any……	Yes	No	Not applicable
33. Problems with your periods	1	2	3
34. If Yes, please indicate the main problem	Irregular		1
	Painful		2
	Heavy		3
	Premenstrual tension		4
	Other		5

Appendix D. Supplementary results for Chapter 3.

One of the differences between the GENDEP and GenPod samples is the frequency with which ADRs were measured were measured on a weekly basis, whilst in GenPod, they were measured at week six and week twelve of the study. In contribute to the discrepant findings between the two studies, the GENDEP analysis was repeated using only the ADR d The significant association between genetic variation in HTR2C and the occurrence of serotonergic ADRs was still difference in the frequency with which ADRs were measured is not responsible for the divergent results in the two samples

Table 10-1: Genetic association with serotonergic adverse drug reactions in GENDEP using only week six and week twelve data (whole MAF= minor allele frequency, n=number of individuals, Obs=total number of observations across 2 timepoints. Odds Ratios are per min significance, when correcting for the number of SNPs testing within the HTR2C gene (p<0.00520)

Gene	C'some	SNP	Location	Allele	MAF	n	Obs	
HTR2C	X	rs6644093	114064023	T/G	0.15	300	503	0.
		rs12846241	113854086	G/T	0.18	300	503	0.
		rs2428700	114010664	A/G	0.14	299	502	0.
		rs4332303	114047867	T/C	0.14	300	503	0.
		rs5946005	114082535	G/A	0.14	300	503	0.
		rs5988087	113934856	T/C	0.16	300	503	0.
		rs11167436	113944060	A/C	0.16	300	503	0.
		rs556677	113822902	A/G	0.16	300	503	0.
		rs543229	113820986	C/T	0.17	300	503	0.

10.1.1 Supplementary analysis of expanded gene region

To consider whether more distant variants were involved in determining individual risk to antidepressant side effects, the region surrounding each candidate gene was expanded to cover 50KB up and downstream. This expanded approach was only possible within the GENDEP sample.

An additional 360 SNPs were identified (the number of new and old SNPs per gene are shown in Table 10-2. Using SNPSpD, the addition of these extra variants was assessed to give a new threshold of significance of $p<7.59x10^{-5}$. These additional variants enabled assessment of variation in the candidate genes *HTR1B*, *HTR1D*, *HTR1F* and *CHRM4*, which was not possible in the primary analysis.

Repeating the analysis in the framework of the primary analysis, each variant was tested for association with cholinergic, serotonergic, adrenergic and histaminergic side effects, using logistic regressions as detailed in the Methods section of Chapter 3. The analyses were conducted in the whole sample, and then in escitalopram and nortriptyline subsets.

The association of rs6644093 with serotonergic ADRs remained significant, exceeding the more stringent significance threshold set for this expanded analysis. Only one of the additional variants showed significant association with any of the four phenotypes (serotonergic, cholinergic, adrenergic and histaminergic ADRs) tested, when performing whole sample and drug-specific analyses. This variant (rs6467694) is 7318bp downstream from the UTR of the CHRM2 gene, and showed significant association with cholinergic outcomes in the whole sample analysis (OR=3.09, 95% CI=1.81-5.29, p=3.75 x 10^{-5}). The function of the variant is unknown, with a minor allele frequency of 9.3% in the GENDEP sample.

Figure 10-1 shows the pattern of association with cholinergic side effects across the *CHRM2* gene. The gene-wide significance threshold for *CHRM2*, when additional variants are included

is p<0.00107; but no other SNPs within the gene reached this threshold. Furthermore, rs6467694 failed to reach gene-wide significance in either of the drug-specific analyses.

The occurrence of cholinergic side effects for each genotype at rs6467694, for the whole sample is shown in Figure 10-2. It should be noted that the prevalence of cholinergic symptoms at baseline (prior to receiving an antidepressant) differs between genotypic categories.

Table 10-2: Additional markers included by extending analysis to encompass 50KB up and downstream of the gene.

Candidate gene	No. of markers in original analysis (1KB)	Additional markers in extended analysis (50KB)	Total markers per gene
SLC6A2	47	14	61
ADRA1A	53	36	89
ADRA1B	14	20	34
ADRA1D	7	27	34
SCL6A4	21	7	28
HTR1A	3	9	12
HTR1B	0	29	29
HTR1D	0	13	13
HTR1E	9	20	29
HTR1F	0	14	14
HTR2A	51	35	86
HTR2B	2	7	9
HTR2C	14	9	23
HTR3A	9	15	24
HTR3B	5	10	15
HTR3C	2	4	6
HTR3D	5	25	30
HTR3E	2	10	12
CHRM1	3	11	14
CHMR2	32	23	55
CHRM3	119	10	129
CHRM4	0	4	4
CHRM5	16	6	22
HRH1	22	2	24
Total no. of SNPs	**436**	**360**	**796**

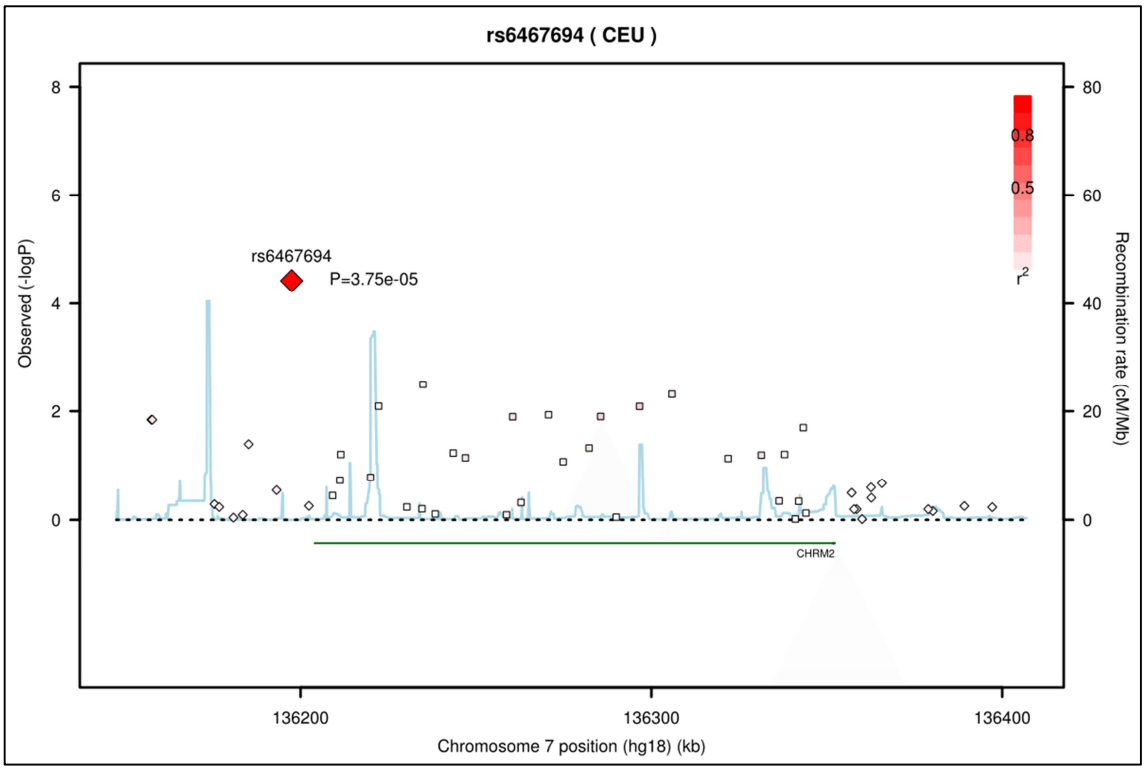

Figure 10-1: Regional association plot showing SNPs within CHRM2 and association with cholinergic side effects (in the whole sample) Generated using SNAP (Johnson et al, 2008)

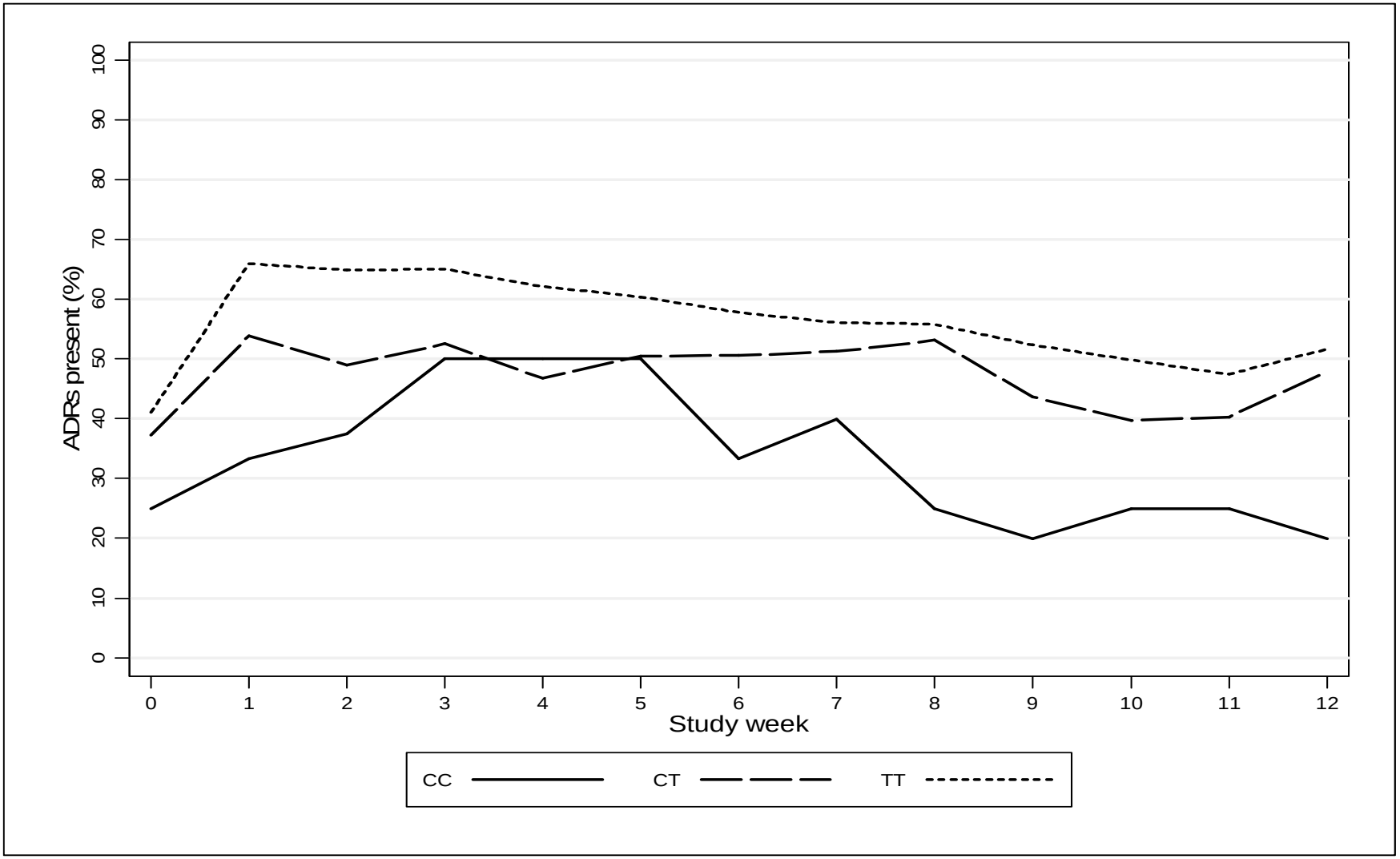

Figure 10-2: Percentage of GENDEP patients reporting cholinergic ADRs, by rs6467694 genotype

Appendix E. Supplementary results for Chapter 5.

Table 10-3: Association between specific ADRs and CYP450 genotype or serum levels of antidepressant, for patients taking nortriptyline

Dry mouth	**n**	**obs**	**p**	**OR**	**SE**
CYP2D6 genotype	251	2182	0.111	0.710	0.153
nortriptyline	184	1888	0.002	1.826	0.362
10-hydroxynortriptyline	180	1845	1.20E-04	2.100	0.405
ratio (10-hydroxynortriptyline: nortriptyline)	178	1830	0.041	1.406	0.234
total (nortriptyline + 10-hydroxynortriptyline)	187	1830	4.97E-05	2.284	0.465
Drowsiness	**n**	**obs**	**p**	**OR**	**SE**
CYP2D6 genotype	251	2178	0.844	1.032	0.164
nortriptyline	184	1884	0.695	1.054	0.142
10-hydroxynortriptyline	180	1841	0.996	0.999	0.123
ratio (10-hydroxynortriptyline: nortriptyline)	178	1826	0.277	1.172	0.171
total (nortriptyline + 10-hydroxynortriptyline)	178	1826	0.794	1.033	0.129
Insomnia	**n**	**obs**	**p**	**OR**	**SE**
CYP2D6 genotype	251	2160	0.395	0.838	0.174
nortriptyline	182	1866	0.486	0.908	0.126
10-hydroxynortriptyline	179	1835	0.655	1.087	0.202
ratio (10-hydroxynortriptyline: nortriptyline)	177	1820	0.086	1.264	0.172
total (nortriptyline + 10-hydroxynortriptyline)	177	1820	0.921	0.985	0.153
Blurred vision	**n**	**obs**	**p**	**OR**	**SE**
CYP2D6 genotype	250	2172	0.681	1.125	0.323
nortriptyline	183	1878	0.849	0.972	0.145
10-hydroxynortriptyline	179	1835	0.514	0.916	0.123
ratio (10-hydroxynortriptyline: nortriptyline)	177	1820	0.775	1.040	0.142
total (nortriptyline + 10-hydroxynortriptyline)	177	1820	0.697	0.953	0.119
Headache	**n**	**obs**	**p**	**OR**	**SE**
CYP2D6 genotype	251	2176	0.127	1.320	0.241
nortriptyline	183	1874	0.023	0.678	0.115
10-hydroxynortriptyline	179	1831	0.325	0.857	0.134
ratio (10-hydroxynortriptyline: nortriptyline)	177	1816	0.220	1.165	0.145
total (nortriptyline + 10-hydroxynortriptyline)	177	1816	0.052	0.713	0.124
Constipation	**n**	**obs**	**p**	**OR**	**SE**
CYP2D6 genotype	251	2181	0.075	0.687	0.145
nortriptyline	184	1887	0.175	1.224	0.183
10-hydroxynortriptyline	180	1845	0.658	0.944	0.122
ratio (10-hydroxynortriptyline: nortriptyline)	178	1830	0.487	0.893	0.146
total (nortriptyline + 10-hydroxynortriptyline)	178	1830	0.472	1.097	0.141
Diarrhoea	**n**	**obs**	**p**	**OR**	**SE**
CYP2D6 genotype	251	2181	0.125	1.898	0.793
nortriptyline	184	1888	0.654	1.113	0.266
10-hydroxynortriptyline	180	1845	0.016	0.652	0.116
ratio (10-hydroxynortriptyline: nortriptyline)	178	1830	0.077	0.624	0.166
total (nortriptyline + 10-hydroxynortriptyline)	178	1830	0.579	0.879	0.204
Increased Appetite	**n**	**obs**	**p**	**OR**	**SE**
CYP2D6 genotype	251	2182	0.978	0.994	0.224
nortriptyline	184	1888	0.191	1.212	0.178

10-hydroxynortriptyline	180	1845	0.768	1.044	0.151
ratio (10-hydroxynortriptyline: nortriptyline)	178	1830	0.751	0.953	0.145
total (nortriptyline + 10-hydroxynortriptyline)	178	1830	0.248	1.174	0.163
Decreased Appetite	**n**	**obs**	**p**	**OR**	**SE**
CYP2D6 genotype	251	2180	0.696	0.923	0.189
nortriptyline	184	1888	0.254	0.828	0.137
10-hydroxynortriptyline	180	1845	0.451	0.871	0.160
ratio (10-hydroxynortriptyline: nortriptyline)	178	1830	0.662	1.086	0.204
total (nortriptyline + 10-hydroxynortriptyline)	178	1830	0.219	0.818	0.134
Nausea/vomiting	**n**	**obs**	**p**	**OR**	**SE**
CYP2D6 genotype	251	2176	0.740	0.911	0.257
nortriptyline	184	1885	0.420	0.849	0.172
10-hydroxynortriptyline	180	1842	0.122	0.763	0.133
ratio (10-hydroxynortriptyline: nortriptyline)	178	1827	0.577	1.090	0.169
total (nortriptyline + 10-hydroxynortriptyline)	178	1827	0.196	0.767	0.157
Problems with urination	**n**	**obs**	**p**	**OR**	**SE**
CYP2D6 genotype	251	2178	0.111	1.591	0.464
nortriptyline	184	1885	0.318	0.704	0.248
10-hydroxynortriptyline	180	1842	0.637	0.923	0.157
ratio (10-hydroxynortriptyline: nortriptyline)	178	1827	0.242	1.233	0.220
total (nortriptyline + 10-hydroxynortriptyline)	178	1827	0.477	0.846	0.199
Problems with sexual function	**n**	**obs**	**p**	**OR**	**SE**
CYP2D6 genotype	247	2087	0.814	0.960	0.166
nortriptyline	181	1812	0.867	0.978	0.130
10-hydroxynortriptyline	177	1769	0.969	0.994	0.150
ratio (10-hydroxynortriptyline: nortriptyline)	175	1754	0.760	0.954	0.148
total (nortriptyline + 10-hydroxynortriptyline)	175	1754	0.958	1.008	0.142
Palpitations	**n**	**obs**	**p**	**OR**	**SE**
CYP2D6 genotype	251	2178	0.933	0.985	0.182
nortriptyline	184	1888	0.775	0.964	0.124
10-hydroxynortriptyline	180	1845	0.567	1.069	0.125
ratio (10-hydroxynortriptyline: nortriptyline)	178	1830	0.096	1.249	0.167
total (nortriptyline + 10-hydroxynortriptyline)	178	1830	0.803	1.031	0.126
Feeling light-headed on standing	**n**	**obs**	**p**	**OR**	**SE**
CYP2D6 genotype	251	2180	0.867	1.033	0.199
nortriptyline	184	1888	0.679	1.055	0.137
10-hydroxynortriptyline	180	1845	0.178	1.187	0.151
ratio (10-hydroxynortriptyline: nortriptyline)	178	1830	0.425	1.106	0.139
total (nortriptyline + 10-hydroxynortriptyline)	178	1830	0.275	1.157	0.155
Feeling like the room is spinning	**n**	**obs**	**p**	**OR**	**SE**
CYP2D6 genotype	250	2167	0.184	1.343	0.298
nortriptyline	183	1873	0.382	0.886	0.122
10-hydroxynortriptyline	179	1830	0.848	1.030	0.159
ratio (10-hydroxynortriptyline: nortriptyline)	177	1815	0.533	1.105	0.177
total (nortriptyline + 10-hydroxynortriptyline)	177	1815	0.622	0.930	0.137
Sweating	**n**	**obs**	**p**	**OR**	**SE**
CYP2D6 genotype	251	2181	0.573	1.128	0.242
nortriptyline	184	1887	0.028	0.720	0.108
10-hydroxynortriptyline	180	1844	0.846	0.974	0.131
ratio (10-hydroxynortriptyline: nortriptyline)	178	1829	0.015	1.424	0.207

total (nortriptyline + 10-hydroxynortriptyline)	178	1829	0.130	0.812	0.112
Increased body temperature	**n**	**obs**	**p**	**OR**	**SE**
CYP2D6 genotype	248	2158	0.367	1.358	0.461
nortriptyline	182	1867	0.007	0.399	0.136
10-hydroxynortriptyline	178	1824	0.403	0.794	0.219
ratio (10-hydroxynortriptyline: nortriptyline)	176	1809	0.098	1.460	0.334
total (nortriptyline + 10-hydroxynortriptyline)	176	1809	0.021	0.514	0.148
Tremor	**n**	**obs**	**p**	**OR**	**SE**
CYP2D6 genotype	251	2181	0.022	0.641	0.124
nortriptyline	184	1886	0.117	1.223	0.157
10-hydroxynortriptyline	180	1843	0.627	1.061	0.130
ratio (10-hydroxynortriptyline: nortriptyline)	178	1828	0.660	1.065	0.154
total (nortriptyline + 10-hydroxynortriptyline)	178	1828	0.146	1.195	0.147
Disorientation	**n**	**obs**	**p**	**OR**	**SE**
CYP2D6 genotype	239	2070	0.643	1.139	0.321
nortriptyline	183	1877	0.879	1.029	0.191
10-hydroxynortriptyline	179	1834	0.604	0.923	0.143
ratio (10-hydroxynortriptyline: nortriptyline)	177	1819	0.837	0.965	0.166
total (nortriptyline + 10-hydroxynortriptyline)	177	1819	0.926	1.018	0.194
Yawning	**n**	**obs**	**p**	**OR**	**SE**
CYP2D6 genotype	251	2181	0.402	0.814	0.200
nortriptyline	184	1887	0.642	0.922	0.161
10-hydroxynortriptyline	180	1844	0.977	0.994	0.205
ratio (10-hydroxynortriptyline: nortriptyline)	178	1829	0.501	1.117	0.184
total (nortriptyline + 10-hydroxynortriptyline)	178	1829	0.708	0.930	0.181
Weight gain	**n**	**obs**	**p**	**OR**	**SE**
CYP2D6 genotype	251	2171	0.563	0.911	0.146
nortriptyline	184	1876	0.361	1.123	0.143
10-hydroxynortriptyline	180	1833	0.464	0.913	0.113
ratio (10-hydroxynortriptyline: nortriptyline)	178	1818	0.454	0.907	0.118
total (nortriptyline + 10-hydroxynortriptyline)	178	1818	0.772	1.038	0.133

Table 10-4: Association between specific ADRs and CYP450 genotype or serum levels of antidepressant, for patients taking escitalopram

Dry mouth	**n**	**obs**	**p**	**OR**	**SE**
CYP2C19 genotype	340	3323	0.964	1.004	0.091
escitalopram	275	2969	6.85E-04	1.480	0.170
desmethylcitalopram	205	2233	0.002	1.420	0.159
ratio (desmethylcitalopram: escitalopram)	204	2221	0.616	0.931	0.133
total (escitalopram + desmethylcitalopram)	204	2221	1.20E-03	1.496	0.186
Drowsiness	**n**	**obs**	**p**	**OR**	**SE**
CYP2C19 genotype	336	3274	0.766	0.978	0.074
escitalopram	272	2930	0.005	1.335	0.137
desmethylcitalopram	203	2206	0.003	1.314	0.122
ratio (desmethylcitalopram: escitalopram)	202	2194	0.293	0.872	0.114
total (escitalopram + desmethylcitalopram)	202	2194	0.007	1.357	0.153
Insomnia	**n**	**obs**	**p**	**OR**	**SE**
CYP2C19 genotype	340	3323	0.936	1.006	0.071
escitalopram	275	2968	0.434	0.919	0.099
desmethylcitalopram	205	2234	9.01E-01	1.014	0.117
ratio (desmethylcitalopram: escitalopram)	204	2222	0.499	1.081	0.125
total (escitalopram + desmethylcitalopram)	204	2222	0.497	0.922	0.110
Blurred vision	**n**	**obs**	**p**	**OR**	**SE**
CYP2C19 genotype	339	3307	0.200	1.123	0.102
escitalopram	275	2963	0.301	1.142	0.147
desmethylcitalopram	205	2227	0.077	1.230	0.144
ratio (desmethylcitalopram: escitalopram)	204	2215	0.381	1.125	0.152
total (escitalopram + desmethylcitalopram)	204	2215	0.237	1.193	0.179
Headache	**n**	**obs**	**p**	**OR**	**SE**
CYP2C19 genotype	339	3311	0.770	0.979	0.070
escitalopram	275	2966	0.983	1.002	0.096
desmethylcitalopram	205	2230	0.669	0.958	0.097
ratio (desmethylcitalopram: escitalopram)	204	2218	0.459	1.085	0.120
total (escitalopram + desmethylcitalopram)	204	2218	0.869	0.982	0.109
Constipation	**n**	**obs**	**p**	**OR**	**SE**
CYP2C19 genotype	340	3322	0.214	0.868	0.099
escitalopram	275	2967	0.477	1.107	0.159
desmethylcitalopram	205	2231	0.131	1.200	0.145
ratio (desmethylcitalopram: escitalopram)	204	2219	0.397	0.865	0.148
total (escitalopram + desmethylcitalopram)	204	2219	0.237	1.196	0.181
Diarrhoea	**n**	**obs**	**p**	**OR**	**SE**
CYP2C19 genotype	339	3311	0.153	0.870	0.085
escitalopram	274	2958	0.032	1.269	0.141
desmethylcitalopram	204	2222	0.608	1.066	0.133
ratio (desmethylcitalopram: escitalopram)	203	2210	4.96E-04	0.597	0.088
total (escitalopram + desmethylcitalopram)	203	2210	0.010	1.368	0.167
Increased Appetite	**n**	**obs**	**p**	**OR**	**SE**
CYP2C19 genotype	338	3299	0.199	0.892	0.080
escitalopram	272	2934	0.815	1.028	0.122
desmethylcitalopram	202	2199	0.522	0.886	0.168
ratio (desmethylcitalopram: escitalopram)	201	2187	0.316	0.835	0.150

total (escitalopram + desmethylcitalopram)	201	2187	0.537	0.905	0.147
Decreased Appetite	**n**	**obs**	**p**	**OR**	**SE**
CYP2C19 genotype	340	3324	0.856	1.016	0.089
escitalopram	274	2960	0.522	0.924	0.113
desmethylcitalopram	204	2224	0.949	0.990	0.151
ratio (desmethylcitalopram: escitalopram)	203	2212	0.544	1.083	0.143
total (escitalopram + desmethylcitalopram)	203	2212	0.940	0.989	0.141
Nausea/vomiting	**n**	**obs**	**p**	**OR**	**SE**
CYP2C19 genotype	336	3276	0.637	1.041	0.089
escitalopram	272	2932	0.452	1.100	0.140
desmethylcitalopram	202	2196	0.174	0.856	0.098
ratio (desmethylcitalopram: escitalopram)	201	2184	0.104	0.806	0.107
total (escitalopram + desmethylcitalopram)	201	2184	0.445	1.104	0.143
Problems with urination	**n**	**obs**	**p**	**OR**	**SE**
CYP2C19 genotype	339	3304	0.977	1.004	0.130
escitalopram	274	2949	0.697	1.078	0.208
desmethylcitalopram	204	2214	0.313	1.221	0.242
ratio (desmethylcitalopram: escitalopram)	203	2202	0.979	0.994	0.225
total (escitalopram + desmethylcitalopram)	203	2202	0.631	1.115	0.253
Problems with sexual function	**n**	**obs**	**p**	**OR**	**SE**
CYP2C19 genotype	337	3220	0.664	1.035	0.082
escitalopram	271	2866	0.200	1.176	0.149
desmethylcitalopram	202	2175	0.468	1.094	0.135
ratio (desmethylcitalopram: escitalopram)	201	2163	0.264	0.859	0.117
total (escitalopram + desmethylcitalopram)	201	2163	0.240	1.167	0.153
Palpitations	**n**	**obs**	**p**	**OR**	**SE**
CYP2C19 genotype	339	3307	0.155	1.116	0.086
escitalopram	274	2953	0.281	1.118	0.116
desmethylcitalopram	204	2217	0.008	1.311	0.133
ratio (desmethylcitalopram: escitalopram)	203	2205	0.278	1.129	0.126
total (escitalopram + desmethylcitalopram)	203	2205	0.298	1.137	0.140
Feeling light-headed on standing	**n**	**obs**	**p**	**OR**	**SE**
CYP2C19 genotype	338	3298	0.120	1.193	0.135
escitalopram	274	2957	0.978	1.004	0.145
desmethylcitalopram	204	2221	0.026	1.276	0.140
ratio (desmethylcitalopram: escitalopram)	203	2209	0.138	1.243	0.182
total (escitalopram + desmethylcitalopram)	203	2209	0.457	1.122	0.173
Feeling like the room is spinning	**n**	**obs**	**p**	**OR**	**SE**
CYP2C19 genotype	338	3297	0.987	1.002	0.096
escitalopram	272	2933	0.160	1.263	0.210
desmethylcitalopram	202	2197	3.28E-05	1.564	0.168
ratio (desmethylcitalopram: escitalopram)	201	2185	0.306	1.225	0.243
total (escitalopram + desmethylcitalopram)	201	2185	0.020	1.523	0.275
Sweating	**n**	**obs**	**p**	**OR**	**SE**
CYP2C19 genotype	339	3311	0.553	1.047	0.082
escitalopram	274	2957	0.474	1.069	0.100
desmethylcitalopram	204	2221	0.658	1.056	0.131
ratio (desmethylcitalopram: escitalopram)	203	2209	0.360	0.901	0.102
total (escitalopram + desmethylcitalopram)	203	2209	0.994	1.001	0.112

Increased body temperature	n	obs	p	OR	SE
CYP2C19 genotype	339	3308	0.808	1.032	0.134
escitalopram	273	2942	0.905	0.983	0.143
desmethylcitalopram	203	2206	0.106	0.723	0.146
ratio (desmethylcitalopram: escitalopram)	202	2194	0.136	0.744	0.147
total (escitalopram + desmethylcitalopram)	202	2194	0.131	0.805	0.116
Tremor	**n**	**obs**	**p**	**OR**	**SE**
CYP2C19 genotype	338	3296	0.812	1.024	0.100
escitalopram	273	2942	0.370	1.138	0.164
desmethylcitalopram	204	2219	0.146	1.215	0.163
ratio (desmethylcitalopram: escitalopram)	203	2207	0.971	1.005	0.143
total (escitalopram + desmethylcitalopram)	203	2207	0.180	1.234	0.194
Disorientation	**n**	**obs**	**p**	**OR**	**SE**
CYP2C19 genotype	337	3286	0.893	0.984	0.115
escitalopram	271	2919	0.540	1.088	0.150
desmethylcitalopram	202	2195	0.253	0.857	0.116
ratio (desmethylcitalopram: escitalopram)	201	2183	0.475	1.199	0.305
total (escitalopram + desmethylcitalopram)	201	2183	0.919	0.981	0.185
Yawning	**n**	**obs**	**p**	**OR**	**SE**
CYP2C19 genotype	338	3300	0.366	1.088	0.101
escitalopram	272	2935	0.154	1.163	0.123
desmethylcitalopram	202	2199	0.205	0.824	0.126
ratio (desmethylcitalopram: escitalopram)	201	2187	0.112	0.804	0.111
total (escitalopram + desmethylcitalopram)	201	2187	0.977	0.997	0.112
Weight gain	**n**	**obs**	**p**	**OR**	**SE**
CYP2C19 genotype	338	3295	0.407	0.923	0.090
escitalopram	273	2941	0.296	0.888	0.101
desmethylcitalopram	203	2206	0.364	0.878	0.126
ratio (desmethylcitalopram: escitalopram)	202	2194	0.571	0.921	0.134
total (escitalopram + desmethylcitalopram)	202	2194	0.132	0.819	0.108

Appendix F. Supplementary results for Chapter 6.

Table 10-5: IPA definitions of functional categories referenced in Chapter 6

IPA functional category	Definition
Haematological Disease	Describes haematological diseases. Some examples include anaemia, leukaemia, haemolysis and thrombosis.
Organismal Injury and Abnormalities	Describes functions associated with injuries and abnormalities of multicellular organisms, primarily humans, mice and rats. This includes abnormal functions such as bleeding, oedema, and haemorrhage of any tissue or organ, as well as functions associated with abnormal tissues such as lesions, ulcers, scars and wounds.
Developmental Disorder	Describes developmental disorders or abnormalities. Some examples include agenesis of organ, hypertrophy of tissue and scoliosis.
Infectious Disease	Describes diseases that result from infection of a multicellular organism, primarily humans, mice and rats, by a pathogenic bacteria, fungus, virus or prion. Some examples include HIV infection, malaria and scrapie.
Organismal Development	Describes functions associated with the normal development of multicellular organisms, primarily humans, mice and rats. This includes those functions directly involved with organismal development, such as the growth and development of organisms, organismal response to stimuli and birth of organisms as well as functions indirectly involved in development such as learning and flow of blood.
Connective Tissue Disorders	Describes diseases of the connective tissue. Some examples include arthritis, rheumatic disease, and developmental disorders of bone.
Cellular Assembly and Organisation	Describes functions associated with subcellular components that are involved in cellular organisation and assembly of cellular substructures. Examples of functions in this category include alignment of actin filaments and aggregation of liposomes.
Nervous System Development and Function	Describes functions associated with the normal development and function of the cells, tissues and organs that make up the nervous system as well as functions specific to the nervous system. Examples include activation of astrocytes and cognition.
Cell Signalling	Describes functions that are involved in intracellular signalling pathways. Specifically it describes functions associated with signalling molecules such as cyclic AMP, nitric oxide, and calcium, signalling functions such as tyrosine phosphorylation and guanine nucleotide exchange, as well as receptor-mediated signalling interactions.
Cancer	Describes functions associated with cancer. This includes any process associated with a tumor, cancer cell or cancerous tissue, as well as any object associated with a cancer process such as transformation and metastasis. This category also includes all cancerous diseases such as lymphoma and neuroblastoma.
Organ Development	Describes functions associated with the normal development, differentiation, and formation of organs. Some examples of these functions include maturation, neurogenesis, function and proliferation of organs.

Table 10-6: Gene probes showing suggestive associations with treatment response in whole sample analysis (p<0.001)

Gene	Probe ID	Corr	P Value	Q Value	Cytoband
HS.242159	ILMN_1822368	0.387	1.159E-05	0.179	
OR52B4	ILMN_1776846	-0.382	1.496E-05	0.179	11p15.4d
PMS2CL	ILMN_2262198	-0.375	2.303E-05	0.179	7p22.1a
CHAF1B	ILMN_1674231	-0.374	2.409E-05	0.179	21q22.12b-q22.13a
KCNH2	ILMN_1739987	0.359	5.210E-05	0.310	7q36.1c
TTC18	ILMN_1784516	-0.354	6.778E-05	0.336	10q22.2a
IQCG	ILMN_1802377	-0.343	1.171E-04	0.455	3q29i
NTRK2	ILMN_2357855	0.331	2.098E-04	0.455	9q21.33a
TRIM69	ILMN_1813430	-0.327	2.493E-04	0.455	15q21.1a
LOC728002	ILMN_3296519	-0.327	2.570E-04	0.455	12p13.31b
MARCH8	ILMN_2341626	0.326	2.596E-04	0.455	10q11.21d
LOC100130451	ILMN_3176545	0.326	2.687E-04	0.455	2q34e
LOC100130624	ILMN_3272476	-0.324	2.852E-04	0.455	22q12.1b
APC	ILMN_1662668	-0.322	3.191E-04	0.455	5q22.2a
ZFP36L1	ILMN_1675448	0.321	3.298E-04	0.455	14q24.1c
CHRFAM7A	ILMN_1791501	-0.321	3.337E-04	0.455	15q13.2a
TMCC2	ILMN_1659024	0.321	3.383E-04	0.455	1q32.1g
ICAM4	ILMN_1681296	0.319	3.585E-04	0.455	19p13.2c
MRPS9	ILMN_1813207	-0.317	3.930E-04	0.455	2q12.1d
LOC650909	ILMN_1657348	-0.316	4.124E-04	0.455	
MIR580	ILMN_3309734	0.313	4.806E-04	0.455	5p13.2c
FAM149A	ILMN_1786278	-0.312	5.045E-04	0.455	4q35.2a
HS.582536	ILMN_1887823	0.310	5.291E-04	0.455	
C3ORF34	ILMN_2288483	0.309	5.571E-04	0.455	3q29g
LOC85390	ILMN_1689294	-0.307	6.057E-04	0.455	11q24.1b
CTLA4	ILMN_2348905	-0.307	6.061E-04	0.455	2q33.2a
STK11	ILMN_1751871	0.307	6.176E-04	0.455	19p13.3i
MKRN1	ILMN_1671583	0.306	6.283E-04	0.455	7q34c
RTN4	ILMN_1693598	0.306	6.314E-04	0.455	2p16.1d
LOC644391	ILMN_1815495	-0.306	6.320E-04	0.455	
CD8B	ILMN_1748601	-0.304	6.895E-04	0.455	2p11.2e
LOC100133435	ILMN_3238148	-0.304	6.972E-04	0.455	
DUSP15	ILMN_1689000	0.304	6.990E-04	0.455	20q11.21b
LOC100132011	ILMN_3248157	0.304	7.034E-04	0.455	17q25.1c
MOXD1	ILMN_1687501	-0.304	7.043E-04	0.455	6q23.2b
OPALIN	ILMN_1666090	0.304	7.086E-04	0.455	10q24.1a
PPP2R2B	ILMN_2298365	0.304	7.141E-04	0.455	5q32d-q32e
DRD5	ILMN_1689043	0.303	7.196E-04	0.455	4p16.1b
RBM38	ILMN_1704079	0.303	7.354E-04	0.455	20q13.31a
KLHL28	ILMN_3251605	0.303	7.396E-04	0.455	14q21.3a-q21.3b
ROGDI	ILMN_1722738	0.302	7.678E-04	0.455	16p13.3b
POFUT1	ILMN_2276758	0.302	7.764E-04	0.455	20q11.21b
CHODL	ILMN_1726575	-0.301	7.920E-04	0.455	21q21.1d
CDCA7	ILMN_1737184	-0.300	8.194E-04	0.455	2q31.1e
HS.133410	ILMN_1819304	-0.300	8.279E-04	0.455	

LIG3	ILMN_2373335	0.300	8.309E-04	0.455	17q12a
LOC100132346	ILMN_3238735	-0.299	8.472E-04	0.455	16q22.3c
ZNF208	ILMN_1662777	0.299	8.567E-04	0.455	19p12c
PTPRM	ILMN_1744937	-0.299	8.595E-04	0.455	18p11.23b-p11.23a
MEF2A	ILMN_3249825	-0.299	8.759E-04	0.455	15q26.3b
RPL6	ILMN_1717490	-0.299	8.782E-04	0.455	12q24.13a
FCAR	ILMN_2279367	0.299	8.794E-04	0.455	19q13.42b
PSMG3	ILMN_1802627	-0.299	8.814E-04	0.455	7p22.3b-p22.3a
FKBP7	ILMN_1717737	-0.297	9.291E-04	0.455	2q31.2b
HS.519298	ILMN_1861556	0.297	9.314E-04	0.455	
FEN1	ILMN_2160929	-0.297	9.422E-04	0.455	11q12.2b
C9ORF37	ILMN_1799320	-0.297	9.430E-04	0.455	9q34.3f
SNORD4B	ILMN_1720794	-0.297	9.535E-04	0.455	17q11.2a

Table 10-7: Gene probes showing suggestive associations with treatment response in escitalopram-specific analysis (p<0.001)

Gene	Probe ID	Corr	P Value	Q Value	Cytoband
IQCG	ILMN_1802377	-0.455	1.721E-05	0.512	3q29i
LOC728905	ILMN_3237395	-0.424	7.193E-05	0.934	1q21.1d
HS.582536	ILMN_1887823	0.414	1.111E-04	0.934	
VWA1	ILMN_1660554	0.403	1.708E-04	0.934	1p36.33a
DRD5	ILMN_1689043	0.398	2.102E-04	0.934	4p16.1b
GNB3	ILMN_2115336	-0.395	2.433E-04	0.934	12p13.31d
PMS2CL	ILMN_2262198	-0.391	2.769E-04	0.934	7p22.1a
SDC1	ILMN_1768953	0.389	3.060E-04	0.934	2p24.1d
ULK4	ILMN_3236866	-0.385	3.559E-04	0.934	3p22.1b
LOC100133435	ILMN_3238148	-0.384	3.626E-04	0.934	
BAG5	ILMN_1728514	-0.384	3.661E-04	0.934	14q32.33a
PSMG3	ILMN_1802627	-0.383	3.765E-04	0.934	7p22.3b-p22.3a
RBMS3	ILMN_1665040	-0.377	4.874E-04	0.999	3p24.1b-p24.1a
LOC100130348	ILMN_3259120	-0.374	5.450E-04	0.999	11q13.1c
LOC100131655	ILMN_3217577	-0.372	5.786E-04	0.999	
OPALIN	ILMN_1666090	0.365	7.571E-04	0.999	10q24.1a
SELS	ILMN_1803744	-0.363	8.023E-04	0.999	15q26.3d
MOXD1	ILMN_1687501	-0.363	8.126E-04	0.999	6q23.2b
RPS6KA2	ILMN_1790801	-0.359	9.346E-04	0.999	6q27c
OR52B4	ILMN_1776846	-0.357	9.766E-04	0.999	11p15.4d

Table 10-8: Gene probes showing suggestive associations with treatment response in nortriptyline-specific analysis (p<0.001)

Gene	Probe ID	Corr	P Value	Q Value	Cytoband
MMP28	ILMN_1752952	-0.683	1.662E-06	0.031	17q12b
KXD1	ILMN_1790951	0.678	2.103E-06	0.031	19p13.11c
SAMD6	ILMN_1691716	0.655	6.065E-06	0.060	9q22.33c
POFUT2	ILMN_2376667	0.606	4.402E-05	0.328	21q22.3e
FLJ41200	ILMN_1694458	0.590	7.662E-05	0.381	9p23a
OR8J3	ILMN_1795000	-0.587	8.596E-05	0.381	11q11c
LOC100133172	ILMN_3286235	0.583	9.676E-05	0.381	8p23.1a
SPTBN2	ILMN_1667079	0.576	1.243E-04	0.381	11q13.1e
LOC644634	ILMN_1776524	0.574	1.311E-04	0.381	1q21.1e
CDIPT	ILMN_1770425	0.571	1.451E-04	0.381	16p11.2d
NVL	ILMN_1712636	-0.570	1.522E-04	0.381	1q42.11b
HS.389313	ILMN_1846771	0.570	1.536E-04	0.381	
DDX19-DDX19L	ILMN_1728069	0.564	1.870E-04	0.428	16q22.1f
CHST2	ILMN_1794011	0.561	2.025E-04	0.431	3q23d
LOC100131735	ILMN_3208233	-0.558	2.226E-04	0.437	
LOC727825	ILMN_1681325	0.549	2.994E-04	0.437	
CTDSP1	ILMN_1728163	0.546	3.222E-04	0.437	2q35e
KLF2	ILMN_1735930	0.544	3.440E-04	0.437	19p13.11f
PLIN2	ILMN_2138765	-0.543	3.542E-04	0.437	9p22.1b
HS.569921	ILMN_1839217	0.543	3.587E-04	0.437	
GMPS	ILMN_3242205	-0.540	3.911E-04	0.437	3q25.31a
HS.242159	ILMN_1822368	0.539	3.944E-04	0.437	
HS.147346	ILMN_1864989	0.539	4.034E-04	0.437	
LOC100130413	ILMN_3265391	0.538	4.057E-04	0.437	
CHAF1B	ILMN_1674231	-0.538	4.117E-04	0.437	21q22.12b-q22.13a
HS.543219	ILMN_1862873	0.537	4.249E-04	0.437	
AKR1A1	ILMN_1774938	0.537	4.253E-04	0.437	1p34.1b
HSFYP1	ILMN_3239912	0.536	4.386E-04	0.437	22q11.1d
LOC390466	ILMN_1665781	-0.533	4.750E-04	0.437	
CLDN6	ILMN_1804531	0.533	4.777E-04	0.437	16p13.3d
CAPN2	ILMN_1716057	-0.533	4.830E-04	0.437	1q41e
UBAC1	ILMN_1807044	0.531	4.984E-04	0.437	9q34.3c
C1ORF120	ILMN_1755835	0.531	5.080E-04	0.437	1q25.3c
MIMT1	ILMN_3237745	-0.528	5.494E-04	0.437	
SNORD50B	ILMN_3247444	-0.525	6.068E-04	0.437	6q14.3c
LOC100128859	ILMN_3258657	-0.524	6.143E-04	0.437	6q13c
PM20D1	ILMN_2293067	0.524	6.256E-04	0.437	1q32.1g
SUV39H2	ILMN_1789351	0.523	6.297E-04	0.437	10p13c
ZNF238	ILMN_1663155	0.522	6.572E-04	0.437	1q44a
LOC100131744	ILMN_3293244	0.522	6.588E-04	0.437	17q12a
KIAA1245	ILMN_1676829	0.520	6.927E-04	0.437	
C4ORF23	ILMN_1704637	-0.520	6.989E-04	0.437	4p16.1c
ETV3	ILMN_1703180	0.519	7.015E-04	0.437	1q23.1b
IGHMBP2	ILMN_1801909	0.518	7.227E-04	0.437	11q13.2b
C2ORF55	ILMN_3243366	0.516	7.775E-04	0.437	2q11.2c
MTRF1L	ILMN_1786684	0.516	7.830E-04	0.437	6q25.2a
CRYBA4	ILMN_1686362	-0.515	7.969E-04	0.437	22q12.1a
PPM1G	ILMN_1806867	-0.515	7.977E-04	0.437	2p23.3a
LOC646786	ILMN_1654915	0.514	8.144E-04	0.437	
C3ORF60	ILMN_1691557	0.514	8.222E-04	0.437	3p21.31d
LOC643993	ILMN_1660393	0.513	8.314E-04	0.437	

CCNY	ILMN_1708991	0.513	8.343E-04	0.437	10p11.21c
LOC642350	ILMN_1762995	0.513	8.427E-04	0.437	9p24.3b
UBE2I	ILMN_1662934	0.513	8.484E-04	0.437	16p13.3e
CHRNA10	ILMN_1776314	0.513	8.492E-04	0.437	11p15.4d
ECHDC1	ILMN_1762134	-0.512	8.667E-04	0.437	6q22.33a
HS.349049	ILMN_1849285	-0.511	8.828E-04	0.437	
RTN2	ILMN_1749115	0.511	8.858E-04	0.437	19q13.32a
HS.121070	ILMN_1903077	0.511	8.976E-04	0.437	

Table 10-9: Table showing top 5 IPA networks identified in each individual variant analysis

Network	Score	Network eligible molecules	Top diseases and functions
Whole sample analysis, individual variants			
1	46	28	Hematological Disease, Organismal Injury and Abnormalities, Developmental Disorder
2	38	24	Cellular Compromise, Cell Death and Survival, Cellular Development
3	37	25	Cell Cycle, Cellular Assembly and Organization, Cancer
4	36	23	Skeletal and Muscular Disorders, Cell Signaling, Post-Translational Modification
5	33	22	Cell Death and Survival, Cell Cycle, DNA Replication, Recombination, and Repair
Escitalopram-specific analysis, individual variants			
1	37	19	Infectious Disease, Organismal Development, Connective Tissue Disorders
2	35	21	Cellular Assembly and Organization, Cellular Function and Maintenance, Cell Cycle
3	30	17	Carbohydrate Metabolism, Lipid Metabolism, Small Molecule Biochemistry
4	26	16	Metabolic Disease, Organismal Injury and Abnormalities, Renal and Urological Disease
5	26	15	Connective Tissue Disorders, Developmental Disorder, Hereditary Disorder
Nortriptyline-specific analysis, individual variants			
1	36	15	Cellular Assembly and Organization, Nervous System Development and Function, Cell Signaling
2	22	10	Digestive System Development and Function, Organ Morphology, Organismal Development
3	21	10	Lipid Metabolism, Small Molecule Biochemistry, Embryonic Development

Table 10-10: Table showing top 5 IPA networks identified using hub genes in the WGCNA analysis

Network	Score	Network eligible molecules	Top diseases and functions
Whole sample analysis, WGCNA			
1	59	23	Cancer, Organ Development, Organismal Injury and Abnormalities
2	33	14	Nervous System Development and Function, Cardiovascular Disease, Organismal Injury and Abnormalities
3	9	5	Cell Signaling, Metabolic Disease, Neurological Disease
4	3	1	Cellular Function and Maintenance, Cell-To-Cell Signaling and Interaction, Hematological System Development and Function
5	2	1	Cellular Development, Cellular Growth and Proliferation, Developmental Disorder
Escitalopram-specific analysis, WGCNA			
1	37	14	Cancer, Organ Development, Organismal Injury and Abnormalities
2	16	7	Antimicrobial Response, Inflammatory Response, Infectious Disease
Nortriptyline-specific analysis, WGCNA			
1	48	20	Cancer, Organ Development, Organismal Injury and Abnormalities
2	44	17	Post-Translational Modification, Hereditary Disorder, Neurological Disease
3	37	19	Lipid Metabolism, Molecular Transport, Small Molecule Biochemistry
4	32	17	Protein Synthesis, Cardiovascular System Development and Function, Connective Tissue Development and Function
5	29	12	Cell Cycle, Embryonic Development, Tissue Morphology

Appendix G. Supplementary results for Chapter 7.

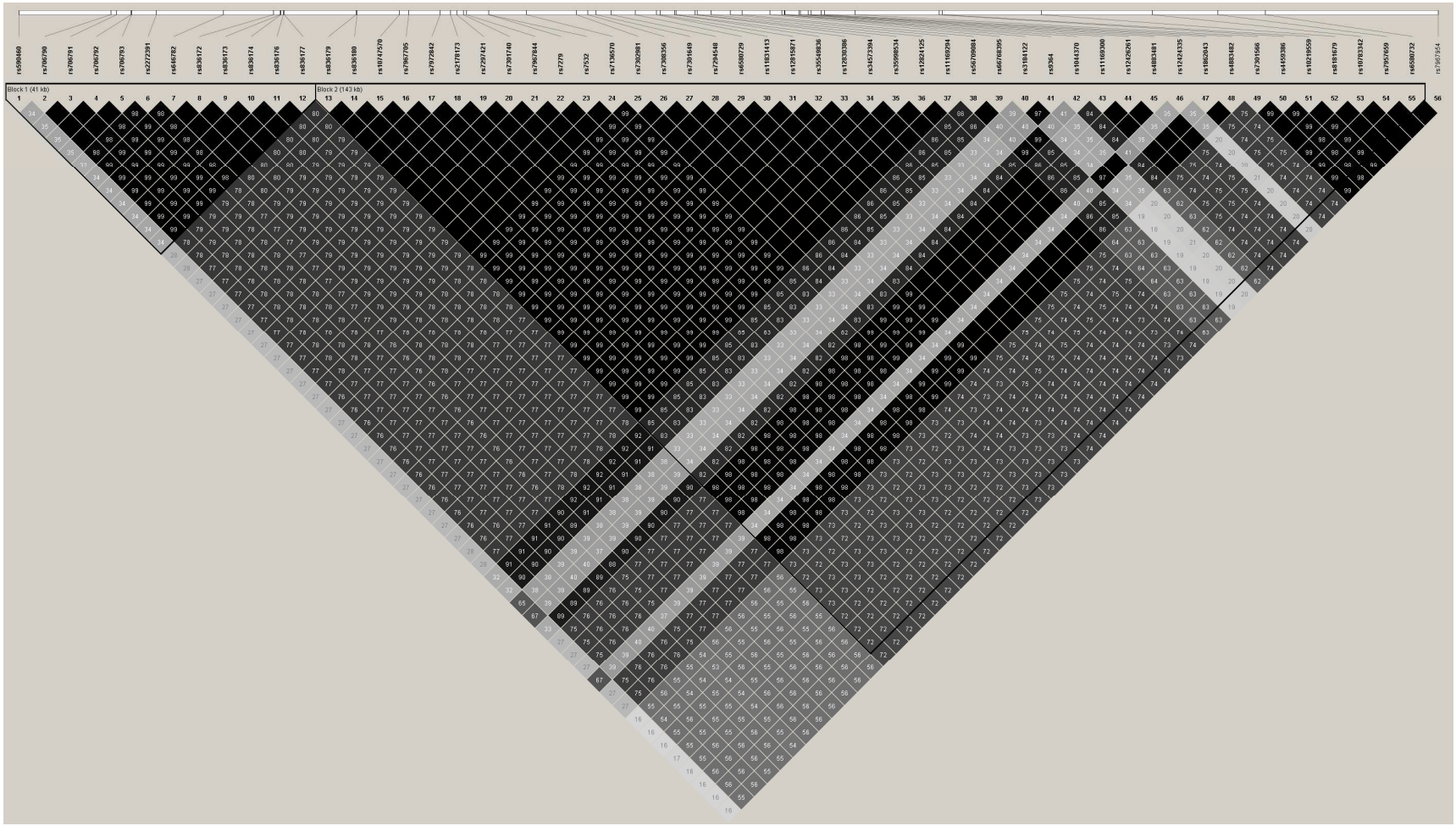

Figure 10-3: Plot showing pattern of LD (as measured using R^2) between SNPs implicated in the CERS5 eQTL locus (drawn using Haploview, Barrett et al, 2005)

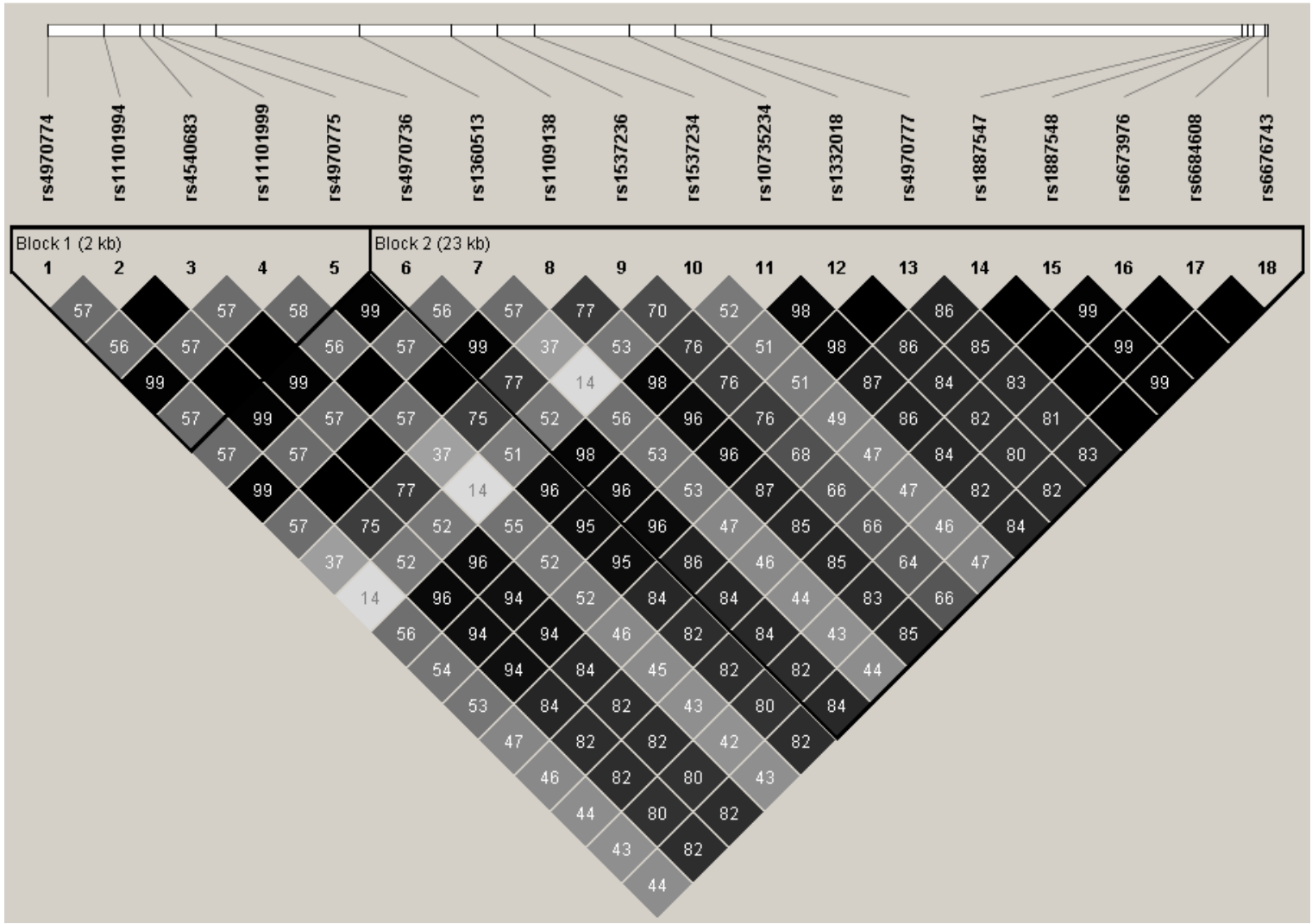

Figure 10-4: Plot showing pattern of LD (as measured using R^2) between SNPs implicated in the GSTM3 eQTL locus (drawn using Haploview, Barrett et al, 2005)